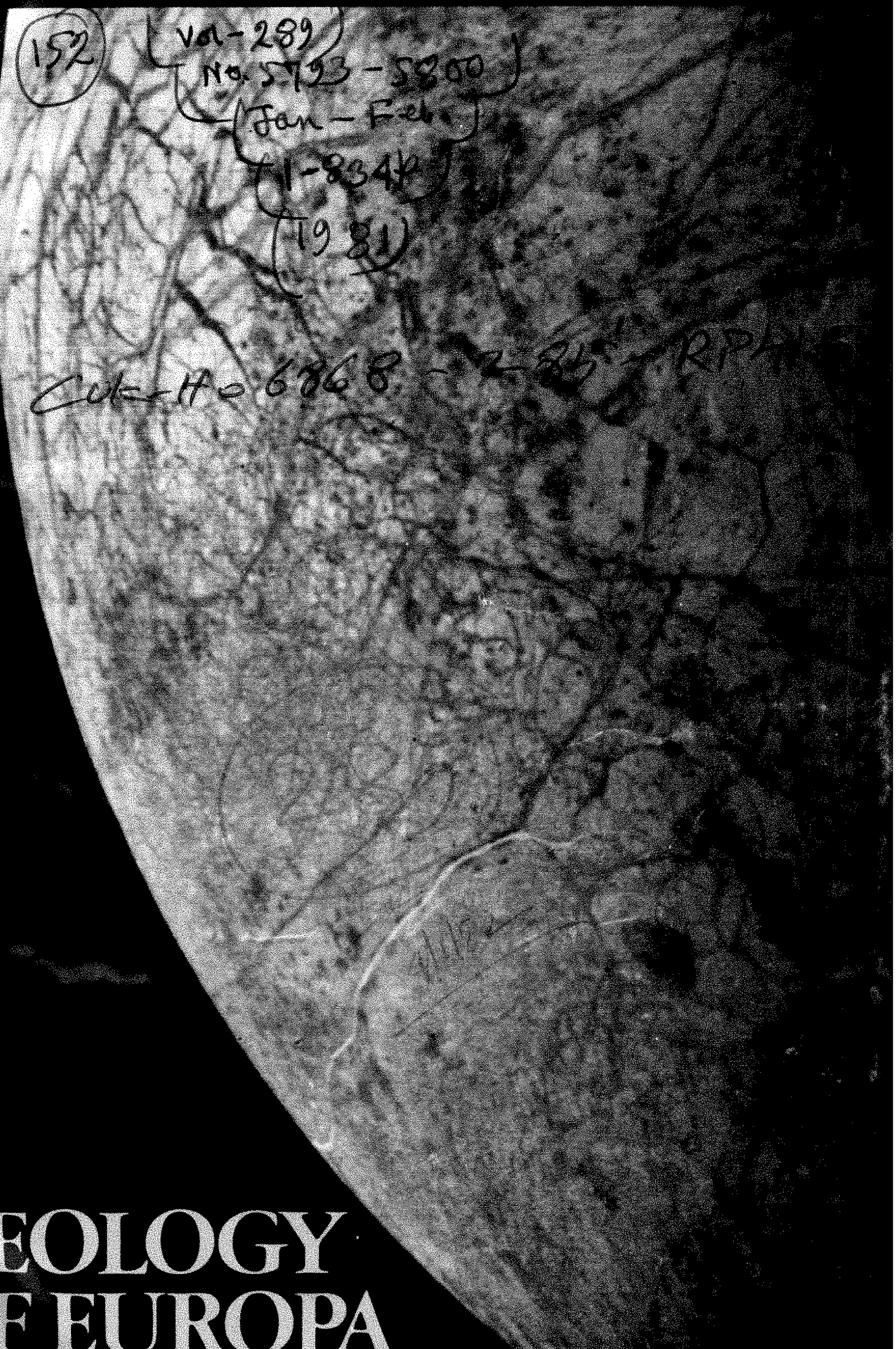
nature

l 289 No 5793 1/8 January 1981 £1.25 $3.00

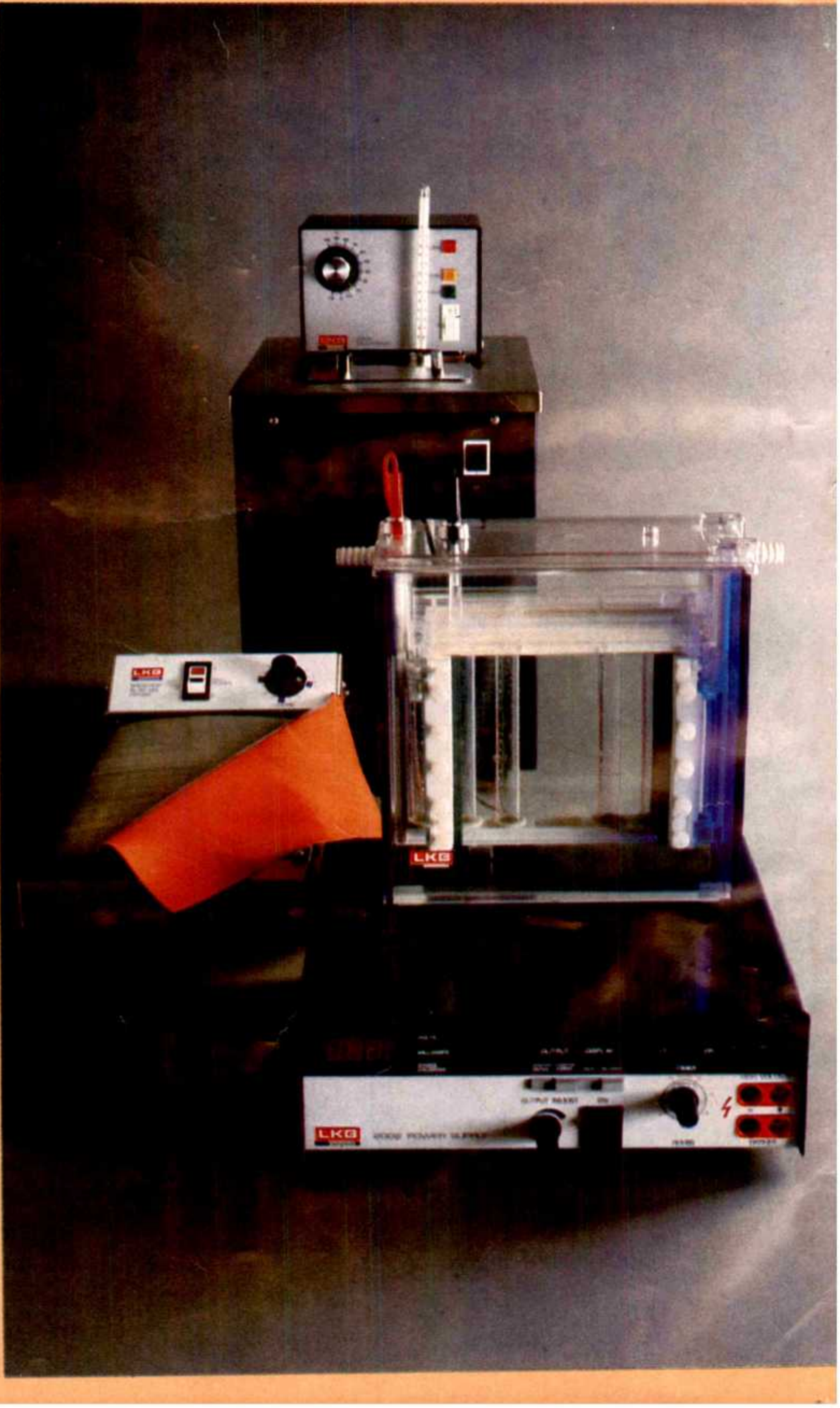


152
Vol-289
No. 5793-5800
Jan-Feb
1-834p
1981
Call-No 6868 - 2-85 - RPL

GEOLOGY OF EUROPA

The new LKB Vertical Electrophoresis System.

Obtain superior electrophoretic separations routinely, with the new vertical electrophoresis instruments from LKB.



LKB 2001 Vertical Electrophoresis Unit
Cast gels without leaking and without messy grease, tape, or agar seals. You can run any of the most widely used polyacrylamide or agarose techniques: SDS, disc, or gradient gels. Perform simple two dimensional separations with our built-in 2-D trough. Circulate buffer, if you wish. Handle heavy workloads with optional twenty well combs and dual slab capability.

LKB 2002 Power Supply
Digital display makes this 500 volt power supply easy to set, and easy to read even from across the lab. A built-in timer assures you reproducible run times and results which are easy to compare.

LKB 2003 Gel Dryer
Preserve your results safely, by drying them onto paper or between cellophane sheets. Accurate temperature control and easy sealing vacuum mean that even hard to handle gradient gels can be dried without cracking.

LKB 2009 MultiTemp Cooling Circulator
Remove the heat produced during electrophoresis to eliminate band skewing and curved migration fronts. The circulating flow of coolant and efficient heat exchanger of the vertical electrophoresis unit permit dependable results, year round, with reduced running times and without wasting water.

Surprisingly low price
You will be pleasantly surprised at how little it costs to acquire the Complete Vertical Electrophoresis System. For full details, call or write to your local LKB office or to the address listed below.

LKB

LKB Instruments Ltd.
232 Addington Road
South Croydon, Surrey CR2 8YD
Telex: 26 44 14 Tel: 657 88 22

Circle No. 16 on Reader Enquiry Card

nature

Volume 289 No.5793 1/8 January 1981

NEWS

Signs of grace (and opposite) for 1981 1

Beating a retreat from Rothschild 2

AAAS plans pressure on education: Dismay at proposed NSF budget cut 3

Slow progress on Greek reforms 3

UK space policy: Year of decision 4

Herbicide safety: Bill of health 4

UK forestry policy: Research wanted 5

Lords look ahead 5

UK clinical research: Future not so rosy 5

New rings for old 5

UK universities: Feeling the pinch 6

Poland freezes Antarctic research 6

Brazil's space programme: Aiming high 6

Nuclear waste: US Senate stalls 7

Correspondence: Badgers and TB (W. Plowright; J. R. Flowerdew); Museum policy (B. A. Wood; D. E. Rosen; D. T. Donovan; M. J. Benton); Halstead replies (L. B. Halstead); Science and values (D. M. Trubshaw); NSF and cryptology (D. N. Langenberg); Behind the fridge (P. A. Jenkins) 8, 105–107

NEWS AND VIEWS

The ocean floor closely observed: Tjeerd H. van Andel challenges the traditional approach to marine geology 9

Did retroviruses evolve from transposable elements: Andrew Flavell comments 10

Magic bullets: Tony Davis reports a recent conference 11

Macroevolution: John Maynard Smith gives a geneticist's view 13

Why is a plant where it is: Peter D. Moore on the individual requirements of plants 14

Galileo would have been pleased: J.E. Guest on the Voyager observations of the Galilean satellites 15

Finding nitrogen fixation genes: John E. Beringer on the genetics of nitrogen fixation 16

ARTICLES

D.C. Pieri — **Lineament and polygon patterns on Europa** 17

G.A. Ransford, A.A. Finnerty and K.D. Collerson — **Europa's petrological thermal history** 21

A.A. Finnerty, G.A. Ransford, D.C. Pieri and K.D. Collerson — **Is Europa surface cracking due to thermal evolution?** 24

E.L. Charnov, R.L. Los-den Hartogh, W.T. Jones and J. van den Assem — **Sex ratio evolution in a variable environment** 27

S. Henikoff, K. Tatchell, B.D. Hall and K.A. Nasmyth — **Isolation of a gene from *Drosophila* by complementation in yeast** 33

S.M. Sun, J.L. Slightom and T.C. Hall — **Intervening sequences in a plant gene — comparison of the partial sequence of cDNA and genomic DNA of French bean phaseolin** 37

285

LETTERS

E.E. Fenimore, W.D. Evans, R.W. Klebesadel, J.G. Laros and J. Terrell — **Spectral evolution of the 5 March 1979 γ burst** 42

H. Ardavan — **Is the light cylinder the site of emission in pulsars?** 44

G.A. Chamberlain, P.J. Cooney and S. Dennison — **Photovoltaic properties of merocyanine solid-state photocells** 45

B.L. Weaver and J. Tarney — **Chemical changes during dyke metamorphism in high-grade basement terrains** 47

B.J. Clough, J.V. Wright and G.P.L. Walker — **An unusual bed of giant pumice in Mexico** 49

D. Bridgwater, J.H. Allaart, J.W. Schopf, C. Klein, M.R. Walter, E.S. Barghoorn, P. Strother, A.H. Knoll and B.E. Gorman — **Microfossil-like objects from the Archaean of Greenland: a cautionary note** 51

B. Nagy, M.H. Engel, J.E. Zumberge, H. Ogino and S.Y. Chang — **Amino acids and hydrocarbons ∿ 3,800-Myr old in the Isua Rocks, southwestern Greenland** 53

Contents continued overleaf

Cover: A computer-enhanced image taken from Voyager 2 of Europa with its surface patterns. A classification scheme for these patterns is presented on page 17 whilst articles on pages 21 and 24 discuss the evolution of the satellite.

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010. ©1981 Macmillan Journals Ltd.

0028-0836/81/01000C-02$01.00

Style that counts

~~R514~~ R415

We've introduced some remarkable features into our new Colony Counter, each aimed at improving an often long and unrewarding but very necessary task.

The compact case styling, simple control layout and improved counter trigger system have resulted in a design which reduces fatigue during prolonged counting operations and requires no special probes or markers.

The counter is simply triggered by depressing the Petri dish. An unobtrusive audible tone increases counting speed by giving an immediate indication of each count action, rendering visual confirmation unnecessary.

Uniform, glare-free illumination of the dish holder minimizes eye strain and clearly defines even small or uncoloured colonies against light or dark backgrounds.

A dishweight sensitivity control is provided to suit each individual's preferred operating pressure and type or size of Petri dish. The unit has three other controls comprising a power on/off switch, a manual reset of the digital display and an audible tone on/off switch.

Circle No. 11 on Reader Enquiry Card.

100
YEARS
SUPPLYING
SCIENCE
1880 · 1980

For further
information and literature
write or telephone:

Gallenkamp
P.O. Box 290,
Technico House,
Christopher Street, London EC2P 2ER
Telephone: 01-247 3211 Telex: 886041

A member of Fisons Scientific Equipment Division

LETTERS

C.H. Miksicek, R.McK. Bird, B. Pickersgill, S. Donaghey, J. Cartwright and N. Hammond **Preclassic lowland maize from Cuello, Belize** 56

R. Lubbock and G.A.B. Shelton **Electrical activity following cellular recognition of self and non-self in a sea anemone** 59

F.S. Walsh and M.A. Ritter **Surface antigen differentiation during human myogenesis in culture** 60

W.A. Siddiqui, S.-C. Kan, K. Kramer, S. Case, K. Palmer and J.F. Niblack **Use of a synthetic adjuvant in an effective vaccination of monkeys against malaria** 64

B. Holtkamp, M. Cramer, H. Lemke and K. Rajewsky **Isolation of a cloned cell line expressing variant $H–2K^k$ using fluorescence-activated cell sorting** 66

J. Robinson, C. Sieff, D. Delia, P.A.W. Edwards and M. Greaves **Expression of cell-surface HLA-DR, HLA-ABC and glycophorin during erythroid differentiation** 68

M. Capron, A. Capron, E.J. Goetzl and K.F. Austen **Tetrapeptides of the eosinophil chemotactic factor of anaphylaxis (ECF-A) enhance eosinophil Fc receptor** 71

A.C. Foster, E.E. Mena, D.T. Monaghan and C.W. Cotman **Synaptic localization of kainic acid binding sites** 73

L.J. Botticelli and R.J. Wurtman **Corticotropin regulates trans-synaptically the activity of septo-hippocampal cholinergic neurones** 75

A.C. Dolphin and P. Greengard **Serotonin stimulates phosphorylation of Protin I in the facial motor nucleus of rat brain** 76

D. Byers, R.L. Davis and J.A. Kiger Jr **Defect in cyclic AMP phosphodiesterase due to the *dunce* mutation of learning in *Drosophila melanogaster*** 79

G. Darby, H.J. Field and S.A. Salisbury **Altered substrate specificity of herpes simplex virus thymidine kinase confers acyclovir-resistance** 81

E. Keshet and Y. Shaul **Terminal direct repeats in a retrovirus-like repeated mouse gene family** 83

G.B. Ruvkun and F.M. Ausubel **A general method for site-directed mutagenesis in prokaryotes** 85

D. Dean, J.L. Yates and M. Nomura ***Escherichia coli* ribosomal protein S8 feedback regulates part of *spc* operon** 89

J. Torbet, J.-M. Freyssinet and G. Hudry-Clergeon **Oriented fibrin gels formed by polymerization in strong magnetic fields** 91

S.B. Brown, A.A. Chabot, E.A. Enderby and A.C.T. North **Orientation of oxygen in oxyhaemoproteins and its implications for haem catabolism** 93

C.R. Woese and R. Gupta **Are archaebacteria merely derived 'prokaryotes'?** 95

MATTERS ARISING

C.P. Wood **Pyroclastic sulphur**
P.W. Francis **Reply to Wood** 97

J.R.S. Fincham **Recombination intermediates**
D.J. Wolgemuth and M.-T. Hsu **Reply to Fincham** 97

C. McGowan and A.J. Baker **Common ancestry for birds and crocodiles?** 97

K.N. Whetstone and L.D. Martin **Reply to McGowan and Baker** 98

BOOK REVIEWS

Dorothy Hodgkin **Sage: A Life of J.D. Bernal (by Maurice Goldsmith)** 99

Peter Lawrence **The Genetics and Biology of Drosophila, Vol. 2d (edited by M. Ashburner and T.R.F. Wright)** 100

G.J. Dockray **Gastrointestinal Hormones (edited by George B. Jerzy Glass)** 100

R.G.C. Bathurst **Early Diagenesis: A Theoretical Approach (by Robert A. Berner)** 101

B.R. Coles **Interactions in Magnetically Ordered Solids (by K.P. Sinha and N. Kumar)** 101

D.E.W. Vaughan **Heterogeneous Catalysis in Practice (by C.N. Satterfield)** 101

R.J. Tayler **Close Binary Stars: Observation and Interpretation (edited by M.J. Plavec and R.K. Ulrich)** 102

John Coles **Guitarrero Cave: Early Man in the Andes (edited by Thomas F. Lynch)** 102

MISCELLANY

Correspondence (continued) 105

Old *Nature*'s almanac for 1981 108

Books received 103

New on the market xv

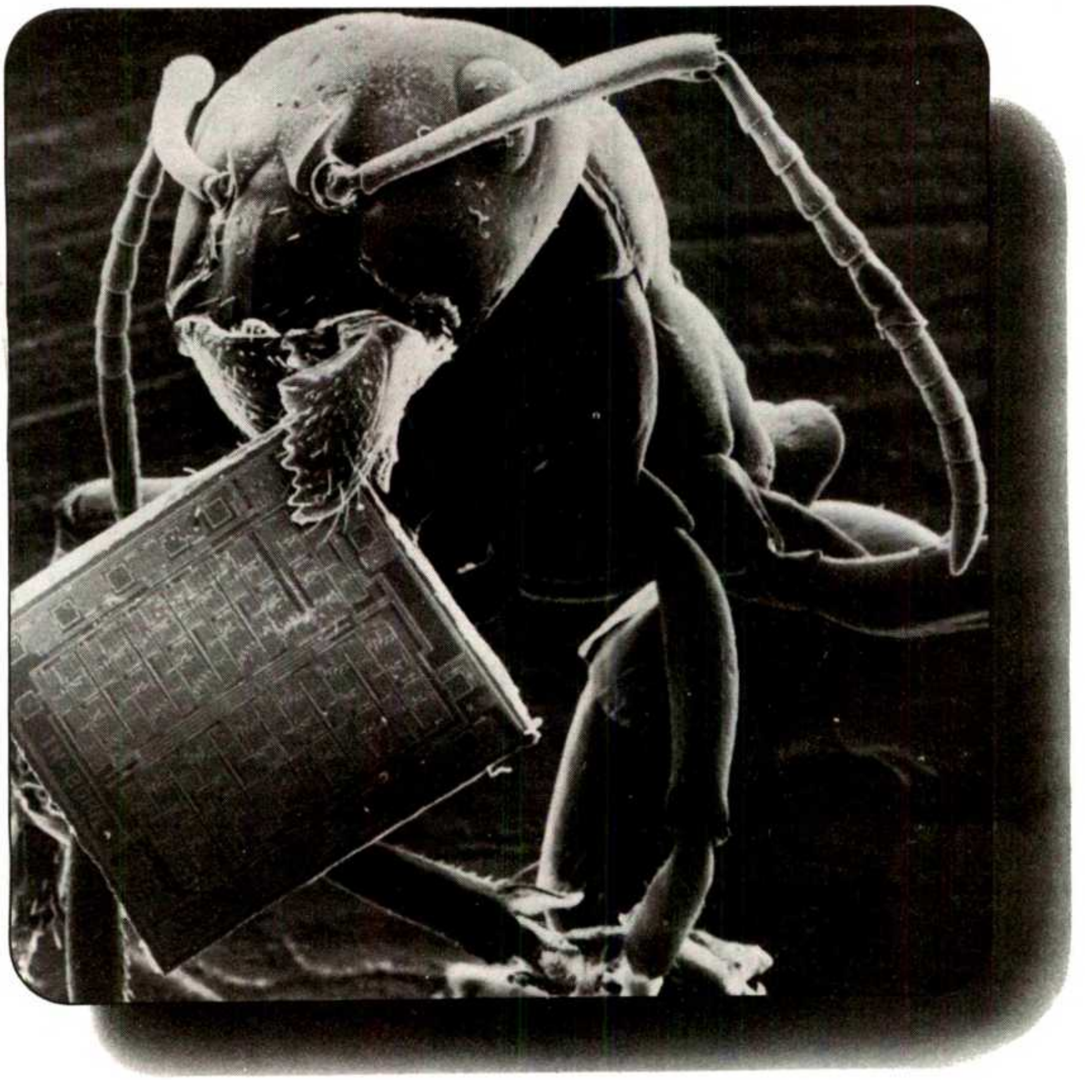
Classified advertising xvii

A no-compromise solution to scanning electron microscopy

The increasing use of scanning electron microscopy in industrial research and product development calls for higher standards and degrees of quantification and improved operational versatility.

Until now however, positional geometry within the specimen environment, along with the many detectors that have to be placed around the specimen, has required an SEM (scanning electron microscope) detector compromise that often produces poor performance in at least one of the operating modes.

So Philips developed a 'no-compromise' scanning electron microscope – the SEM 505 – in which a revolutionary, fibre-optics based, Multi-Function Detector (MFD) interfaces the detecting elements with a photomultiplier/pre-amplifier assembly that can be placed outside of the specimen chamber; thus permitting high flexibility in detector placement. The result? Optimum detection performance in all modes because detection geometry is no longer limited by fixed construction or precise port location.



Invitation

Please visit us at **Microelectronics** – international conference and exhibition of microelectronic production and test equipment – in **Eindhoven, Holland.** February 4,5 and 6, 1981.

PHILIPS

nature

Volume 289 No.5794 15 January 1981

NEWS

Too soon to build laser weapons systems 109

What if Britain wants not to change? 110

Reagan assailed by laser weapons lobbies: Hawks say go: Doves argue caution 111

European science portfolios still for grabs 111

Nucleotide sequences: Too many banks? 112

Polish pollution: Smelter shuts down 112

Human experimentation: Indian rules 113

Vitamin deficiencies: Rickets at bay 113

Research development: Soliciting custom 113

British technology: New direction 114

Latin America: Freedom lost 114

Correspondence: Explosive reaction (A. De Volpi); Role for ESF? (M. K. Wallis); Theories of cancer (S. S. Epstein); Spanish science (B. A. Bolt) 115

NEWS AND VIEWS

Strife over visual cortical function: D.M. MacKay asks if 'feature detector' and 'spatial frequency' models are real alternatives. 117

Yeast genetics and molecular biology: Terrance G. Cooper reports a recent international conference. 119

Remote sensing of the oceans: D.J. Webb describes applications of centimetre radar. 120

Multiple methylation and bacterial adaptation: Judith Armitage looks at a sensory transduction model system. 121

Photosynthesis in the Greek sunshine: B. Halliwell and M.C.W. Evans report the fifth international photosynthesis congress. 122

Interstellar media in other galaxies: A correspondent reports a recent workshop. 123

Precambrian orogenies and polar wandering: J.C. Briden looks at plate movement in the earliest geological times. 125

REVIEW ARTICLE

S.S. Epstein and J.B. Swartz **Fallacies of lifestyle cancer theories** 127

ARTICLES

T.C. Onstott and R.B. Hargraves **Proterozoic transcurrent tectonics: palaeomagnetic evidence from Venezuela and Africa** 131

W.D. Smith **The *b*-value as an earthquake precursor** 136

M.F. Mescher, M.J.L. Jose and S.P. Balk **Actin-containing matrix associated with the plasma membrane of murine tumour and lymphoid cells** 139

H.L. Klein and T.D. Petes **Intrachromosomal gene conversion in yeast** 144

A. Shimizu, N. Takahashi, Y. Yamawaki-Kataoka, Y. Nishida, T. Kataoka and T. Honjo **Ordering of mouse immunoglobulin heavy chain genes by molecular cloning** 149

LETTERS

J.S. Miller, R.R.J. Antonucci and W.C. Keel **Variability of the double quasar 0957+561 A, B** 153

P.E. Johnson, D.W. Smith and R.W. Shorthill **An outburst of comet Tempel-2 observed spectrophotometrically** 155

L.A. Bursill, J.M. Thomas and K.J. Rao **Stability of zeolites under electron irradiation and imaging of heavy cations in silicates** 157

E. Borgarello, J. Kiwi, E. Pelizzetti, M. Visca and M. Grätzel **Photochemical cleavage of water by photocatalysis** 158

R.W. Carlson, G.W. Lugmair and J.D. Macdougall **Crustal influence in the generation of continental flood basalts** 160

D.H. Keen, R.S. Harmon and J.T. Andrews **U series and amino acid dates from Jersey** 162

Contents continued overleaf

Cover: An artificial lichen consisting of the mycobiont *Cladonia cristatella* and the alga *Trebouxia erici*, here showing envelopment by the mycobiont. Experiments discussed on page 169 suggest that the relationship is one of controlled parasitism.

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/02000C-02$01.00

Raven Press

Journal of Cerebral Blood Flow and Metabolism

Official Journal of the International Society of Cerebral Blood Flow and Metabolism

Editor-in-Chief
A. Murray Harper, M.D.
University of Glasgow
Wellcome Surgical Institute
Garscube Estate
Bearsden Road
Bearsden, Glasgow G61 1QH,
Scotland

Assistant Editor:
J. McCulloch, *Glasgow*

Deputy Chief Editors:
K.-A. Hossman, *Cologne*
M. E. Raichle, *St. Louis*
B. K. Siesjö, *Lund*
L. Sokoloff, *Bethesda*

Editorial Board:
A. Bès, *Toulouse*
E. Betz, *Tübingen*
B. Chance, *Philadelphia*
R. C. Collins, *St. Louis*
C. Crone, *Copenhagen*
T. E. Duffy, *New York*
W. Feindel, *Montreal*
J. D. Fenstermacher, *Bethesda*
J. A. Ferrendelli, *St. Louis*
C. Fieschi, *Rome*
F. Gotoh, *Tokyo*
D. D. Heistad, *Iowa City*
D. H. Ingvar, *Lund*
C. Kennedy, *Bethesda*
I. Klatzo, *Bethesda*
N. A. Lassen, *Copenhagen*
D.-W. Lübbers, *Dortmund*
D. G. MacDowall, *Leeds*
E. T. MacKenzie, *Paris*
L. Mela, *Michigan*
W. D. Obrist, *Philadelphia*
Y. Olsson, *Uppsala*
Ch. Owman, *Lund*
H. Pappius, *Montreal*
C. S. Patlak, *Bethesda*
M. E. Phelps, *Los Angeles*
F. Plum, *New York*
S. I. Rapaport, *Bethesda*
M. Reivich, *Philadelphia*
P. Scheinberg, *Miami*
L. Symon, *London*

Honorary Editor:
Seymour S. Kety, *Cambridge*

Aims and Scope:

This quarterly journal will be the official publication of the newly formed International Society of Cerebral Blood Flow and Metabolism. It will gather, in one publication, new information on experimental, theoretical, and clinical aspects of brain circulation and metabolism. It will be of relevance to neurologists, neurochemists, physiologists, pharmacologists, anesthesiologists, neuroradiologists, neuropathologists and neuroscientists in general. Papers will be reviewed by a distinguished editorial board and carefully chosen experts in various fields. Reviewing will be rigorous in order to attain the highest standards of scholarship.

Features:

The following types of material will appear on a regular basis:

- *Original articles* on all fields relating to cerebral blood flow and metabolism including experimental physiological, pharmacological, and biochemical studies, methodological and theoretical papers, and clinical reports.
- *Short communications* reporting concise pieces of work that lead to positive conclusions. These communications should not normally exceed 1500 words and should not contain more than 2 figures.
- *Reviews* on important aspects from invited contributors.
- *Editorials* on subjects of a timely nature.
- Announcements of scientific meetings related to the field.
- Book reviews of new books of interest to the readership.

Circle No.04 on Reader Enquiry Card.

Call for Papers:

The submission of suitable articles for editorial review is welcomed. Articles must be in English and three copies should be submitted. References in the text should be cited by names of authors and year of publication. They should be listed in alphabetical order at the end of the text. Journal titles should be abbreviated in accordance with Index Medicus. A detailed list of instructions to contributors can be obtained from the Publisher or the Editor-in-Chief. The format of the journal will permit high quality reproduction of plates with minimal reduction. Articles should be sent directly to the Editor-in-Chief: Dr. A. M. Harper, M.D., University of Glasgow, Wellcome Surgical Institute, Garscube Estate, Bearsden Road, Bearsden, Glasgow G61 1QH, Scotland.

Subscription Order Form

Date________

To: **RAVEN PRESS**
1140 Avenue of the Americas
New York, N.Y. 10036, U.S.A.

Please enter my subscription for Volume 1 (1981) of **Journal of Cerebral Blood Flow and Metabolism**

Published quarterly.
Personal subscription rates:
U.S. $85.00; elsewhere $95.00.
Institutional subscription rates:
U.S. $95.00; elsewhere $105.00.
Air delivery is *included* for European and Mediterranean countries; for air service *elsewhere* add $15.00.

Name (please print)________________
Address________________
____
____
Country________________
Postal Code________________

LETTERS

J.T. Andrews, P.T. Davis, W.N. Mode, H. Nichols and S.K. Short	**Relative departures in July temperatures in northern Canada for the past 6,000 yr**	**164**
A.K. Behrensmeyer and L.F. Laporte	**Footprints of a Pleistocene hominid in northern Kenya**	**167**
V. Ahmadjian and J.B. Jacobs	**Relationship between fungus and alga in the lichen *Cladonia cristatella* Tuck**	**169**
P.G. Falkowski and Z. Dubinsky	**Light–shade adaptation of *Stylophora pistillata*, a hermatypic coral from the Gulf of Eilat**	**172**
D.A.S. Smith	**Heterozygous advantage expressed through sexual selection in a polymorphic African butterfly**	**174**
G.B. Quistad, D.C. Cerf, D.A. Schooley and G.B. Staal	**Fluoromevalonate acts as an inhibitor of insect juvenile hormone biosynthesis**	**176**
N. Haga and K. Hiwatashi	**A protein called immaturin controlling sexual immaturity in *Paramecium***	**177**
B.J. Mathieson, S.O. Sharrow, Y. Rosenberg and U. Hämmerling	**Lyt 1^+23^- cells appear in the thymus before Lyt 123^+ cells**	**179**
U.H. Koszinowski and M. Kramer	**Selective inhibition of T suppressor-cell function by a monosaccharide**	**181**
B. Fagg	**Is erythropoietin the only factor which regulates late erythroid differentiation?**	**184**
M. Ennis, A. Truneh, J.R. White and F.L. Pearce	**Inhibition of histamine secretion from mast cells**	**186**
H.G. Muchmore, A.J. Parkinson, J.E. Humphries, E.N. Scott, D.A. McIntosh, L.V. Scott, M.K. Cooney and J.A.R. Miles	**Persistent parainfluenza virus shedding during isolation at the South Pole**	**187**
J.S.M. Peiris, S. Gordon, J.C. Unkeless and J.S. Porterfield	**Monoclonal anti-Fc receptor IgG blocks antibody enhancement of viral replication in macrophages**	**189**
P.-F. Lin and F.H. Ruddle	**Murine DNA repair gene located on chromosome 4**	**191**
J. Peters and J. Jagger	**Inducible repair of near-UV radiation lethal damage in *E. coli***	**194**
M. Fogliano and P.F. Schendel	**Evidence for the inducibility of the *uvrB* operon**	**196**
A. Levy and M. Noll	**Chromatin fine structure of active and repressed genes**	**198**
A.S. Benight, R.M. Wartell and D.K. Howell	**Theory agrees with experimental thermal denaturation of short DNA restriction fragments**	**203**
L.J. Eaves and A.C. Heath	**Detection of the effects of asymmetric assortative mating**	**205**

BOOK REVIEWS

Mary Hesse	**The Scientific Image (by Bas C. van Fraassen)**	**207**
A.F. Harding	**A Dictionary of Terms and Techniques in Archaeology (by Sara Champion)**	**208**
Hugh Thomas	**The Human Condition (by W.H. McNeill)**	**208**
C.R.W. Spedding	**Biochemical and Photosynthetic Aspects of Energy Production (edited by Anthony San Pietro)**	**209**
Anthony E. Pegg and Hannu Pösö	**Polyamines in Biomedical Research (edited by Joseph M. Gaugas)**	**209**
Peter S. Liss	**Air in Danger: Ecological Perspectives of the Atmosphere (by G. Breuer)**	**210**
A.J. McCaffery	**Introduction to Molecular Energy Transfer (by James T. Yardley)**	**210**

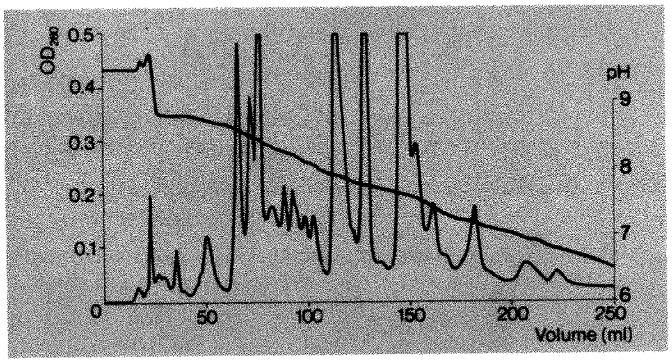
MISCELLANY

100 years ago	**126**
Classified advertising	**xxv**

PRODUCT REVIEW This week's product review (page xix) covers chromatography.

CHROMATOFOCUSING. A NEW HIGH RESOLUTION TECHNIQUE FOR PROTEIN SEPARATION.

If you think chromatofocusing sounds like an interesting combination of two well known techniques, you're right. But if you think its probably only half as good as either of them you're wrong. Chromatofocusing gives you the best of both.



Separation of proteins from mammalian muscle by chromatofocusing.

Separations according to pI

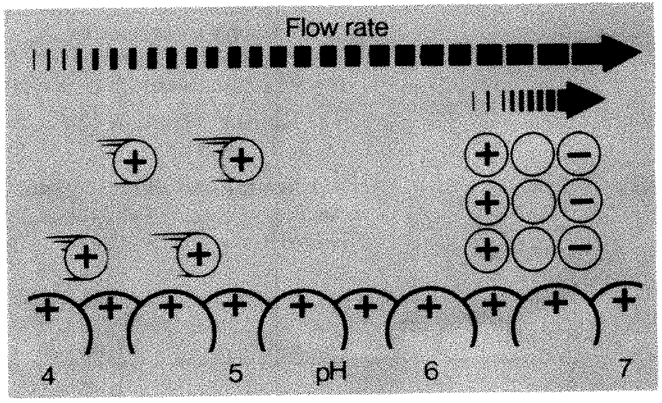
As you've probably guessed, chromatofocusing separates proteins according to isoelectric point. All you need is a column, and the new products Polybuffer™ and Polybuffer exchangers. All you do is equilibrate and pack your column with Polybuffer exchanger, and elute at a lower pH with Polybuffer.

High resolution, high capacity

What you get is a highly selective preparative method for purifying proteins. In addition, focusing effects ensure that samples elute as sharp zones.

Column chromatography with focusing effects

When protein is applied to a column of Polybuffer exchanger below its pI, it migrates downwards in the pH gradient until pH = pI. Beyond this point it binds to the gel. However, as more Polybuffer is added to the column, the pH decreases and the protein elutes according to its pI. Zones are focused since molecules below their pI are repelled from the gel and so catch up with the leading edge of the zone.



Self-generated pH gradients

Chromatofocusing works because Polybuffer exchangers are substituted with basic buffering groups. When these groups are titrated with Polybuffer, at a lower pH, a linear pH gradient forms automatically in the column.

Easy to use Chromatofocusing kit

Two gels and three buffers cover the range pH 11–4. If you're starting from scratch, you should order the Chromatofocusing kit (order no. 17-0520-01), which contains full instructions for use and sufficient material to cover the range pH 9-4.

More information is available in the booklet "Chromatofocusing with Polybuffer and Polybuffer exchangers" so remember to request it when you place an order for the Chromatofocusing kit. We'll accept your order by phone or by post. Once you've tried chromatofocusing you'll wonder how you managed without it.

Pharmacia Fine Chemicals AB
Box 175
S-751 04 Uppsala 1
Sweden

Pharmacia
Fine Chemicals

Circle No. 25 on Reader Enquiry Card.

nature

Volume 289 No.5795 22 January 1981

NEWS

Towards what shining city, which hill? 211
Responsibility for trust in research 211
More pressures on Rothschild system: Geology project threatened by more cuts 213
Carter's last budget asks for more 213
Promises for President Reagan to deny 214
Argentinian power: Soviets help 214
Large electron project: Swedish cloud 215
Artificial hormones in meat: European register? 215
Soviet chemical industry: Effective economy 215
European research: Counting costs 216
Swedish research: Lucky science 216
Graduate education: Too many laggards 217
British universities: More confusion 217
Correspondence: Badgers guilty (Frank Yates); Rothschild retreat (R.R. Porter); Creating problems (Thomas H. Jukes); Genes and race (Steven Rose); Energetic consensus (F.H. Malpress) 218, 335

NEWS AND VIEWS

Deep-ocean hydrothermal vent communities: J.T. Enright *et al.* discuss food sources for the vent animals 219
A new stingray from South Africa: Alwyne Wheeler reports a recent discovery 221
Untangling ataxia-telangiectasia: Bryn A. Bridges and David G. Harnden discuss this genetic disease 222
The spruce invasion: Peter D. Moore traces the expansion of the spruce tree since the last glaciation 223
New vistas in soil research: Nuclear magnetic resonance is the technique of the future according to Hamish Anderson 323
Inclusion compounds: J.E.D. Davies reports an international symposium on crystal complexes 224
The source of meteorites in time and space: Robert Hutchison asks from where in the Solar System meteorites originate 225
Holographic bubble chambers: David J. Miller looks at the technology for studying fundamental particles 226

A case of cell line contanimation

D. Dickson **Contaminated cell lines** 227
N.L. Harris, D.L. Gang, S.C. Quay, S. Poppema, P.C. Zamecnik, W.A. Nelson-Rees and S.J.O'Brien **Contaminations of Hodgkin's disease cell cultures** 228

ARTICLES

M. Wyss, A.C. Johnston and F.W. Klein **Multiple asperity model for earthquake prediction** 231
J.H. Leven, I. Jackson and A.E. Ringwood **Upper mantle seismic anisotropy and lithospheric decoupling** 234
A.J.S. Klar, J.N. Strathern, J.R. Broach and J.B. Hicks **Regulation of transcription in expressed and unexpressed mating type cassettes of yeast** 239
K.A. Nasmyth, K. Tatchell, B.D. Hall, C. Astell and M. Smith **A position effect in the control of transcription at yeast mating type loci** 244
G.F. Sprague Jr, J. Rine and I. Herskowitz **Homology and non-homology at the yeast mating type locus** 250
J.E. Majors and H.E. Varmus **Nucleotide sequences at host–proviral junctions for mouse mammary tumour virus** 253
C.Van Beveren, J.A. Galleshaw, V. Jonas, A.J.M. Berns, R.F. Doolittle, D.J. Donoghue and I.M. Verma **Nucleotide sequence and formation of the transforming gene of a mouse sarcoma virus** 258
J.W. Weisel, G.N. Phillips Jr and C. Cohen **A model from electron microscopy for the molecular structure of fibrinogen and fibrin** 263

LETTERS

S.A. Stephens **Cosmic ray antiprotons and modified closed galaxy model** 267
W.-H. Ip **Expanding haloes in cometary comae** 269

Contents continued overleaf

Cover: A 463 million year old meteorite discovered in Ordovician limestone. Extending to the left is a fossilised nautiloid cephalopod which seems to have been struck and killed by the meteorite!
This issue contains the author index to Volume 288.

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/04000C-02$01.00

Spiders

E.O. Wilson: Let me remind you that people everywhere, a large percent of the population, at a very early age have already developed a deep horror at the sight of snakes or spiders with nothing more than gentle nudging from their parents, if that. Yet, in spite of the fact that parents constantly reinforce their children against going near electric sockets, automobiles, knives and the like, phobias against such objects rarely develop.

Marvin Harris: Let's go back again to the possibility that these phobias are genetically programmed – which I'm willing to grant. The overwhelming bulk of the socially conditioned response repertories of different human societies consists, by your own admission, of culturally determined rather than genetically determined traits. Then it seems to me that when one offers a cogent culturological explanation of these
has to be considered that
offered in

from a debate between E.O. Wilson and Marvin Harris
n sociobiology. In The Sciences. From family
onstellations to galaxies, The Sciences takes the critical
ook. Handsome, original, lively, The
Sciences represents the best science
commentary you'll find anywhere on
human behavior, culture, the physical
sciences — by Ashley Montagu,
Stephen Jay Gould, B.F. Skinner,
Napoleon Chagnon, Peter Medawar
... Subscribe today.
It's only $12.50 a year.

The Sciences. More than science.

☐ Here's my check or money order for $12.50 ($2.50 off the newsstand price) for the next ten issues of **The Sciences.**

Foreign individuals add $5.00
Foreign institutions add $10.00
US institutions add $5.00

☐ Bill me later.

Name

Address

City

State

Zip

The Sciences
2 East 63rd Street
New York, N.Y. 10021

NAT 7

Published by The New York Academy of Sciences

ances of human
such gigantic phenomena as the
gence of state-organized societies or the
emergence of stratified classes and castes. The enormous differences between industrial civilization and

Circle No. 03 on
Reader Enquiry Card.

C.R.V. Raman and Y.P. Rao **Blocking highs over Asia and monsoon droughts over India** 271

E. Murad, W. Swider and S.W. Benson **Possible role for metals in stratospheric chlorine chemistry** 273

P.F. Barron and M.A. Wilson **Humic soil and coal structure study with magic-angle spinning ^{13}C CP-NMR** 275

R.F. Livaccari, K. Burke and A.M.C. Şengör **Was the Laramide orogeny related to subduction of an oceanic plateau?** 276

C.J. Hawkesworth, J.D. Kramers and R. McG. Miller **Old model Nd ages in Namibian Pan-African rocks** 278

M.F. Sheridan, F. Barberi, M. Rosi and R. Santacroce **A model for Plinian eruptions of Vesuvius** 282

P. Thorslund and F.E. Wickman **Middle Ordovician chondrite in fossiliferous limestone from Brunflo, central Sweden** 285

T.E. DeLaca, D.M. Karl and J.H. Lipps **Direct use of dissolved organic carbon by agglutinated benthic foraminifera** 287

C.I. Rose and D.L. Hawksworth **Lichen recolonization in London's cleaner air** 289

E. Martinoia, U. Heck and A. Wiemken **Vacuoles as storage compartments for nitrate in barley leaves** 292

D. Edgar, Y.-A. Barde and H. Thoenen **Subpopulations of cultured chick sympathetic neurones differ in their requirements for survival factors** 294

C. Tickle **The number of polarizing region cells required to specify additional digits in the developing chick wing** 295

D.E. Lewis, J.V. Giorgi and N.L. Warner **Flow cytometry analysis of T cells and continuous T-cell lines from autoimmune MRL/l mice** 298

A.G. Bird, S.M. McLachlan and S. Britton **Cyclosporin A promotes spontaneous outgrowth *in vitro* of Epstein–Barr virus-induced B-cell lines** 300

L.H. Perrin, E. Ramirez, P.H. Lambert and P.A. Miescher **Inhibition of *P. falciparum* growth in human erythrocytes by monoclonal antibodies** 301

I. Vlodavsky and D. Gospodarowicz **Respective roles of laminin and fibronectin in adhesion of human carcinoma and sarcoma cells** 304

J. Michaelson, A. Ferreira and V. Nussenzweig ***cis*-Interacting genes in the *S* region of the murine major histocompatibility complex** 306

A. Juretić, Z.A. Nagy and J. Klein **Detection of CML determinants associated with *H-2* controlled E_β and E_α chains** 308

R.L. Trelstad, K.R. Lawley and L.B. Holmes **Nonenzymatic hydroxylations of proline and lysine by reduced oxygen derivatives** 310

H. Chap, G. Mauco, M.F. Simon, J. Benveniste and L. Douste-Blazy **Biosynthetic labelling of platelet activating factor from radioactive acetate by stimulated platelets** 312

M.T. Nelson and M.P. Blaustein **Effects of ATP and vanadate on calcium efflux from barnacle muscle fibres** 314

B.G. Livett, D.M. Dean, L.G. Whelan, S. Udenfriend and J. Rossier **Co-release of enkephalin and catecholamines from cultured adrenal chromaffin cells** 317

B.J. Wisnieski and J.S. Bramhall **Photolabelling of cholera toxin subunits during membrane penetration** 319

C. DeLisi **The magnitude of signal amplification by ligand-induced receptor clustering** 322

S. Haber, M. Ikegami, N.B. Bajet and R.M. Goodman **Evidence for a divided genome in bean golden mosaic virus, a geminivirus** 324

R. Diaz, K. Nordström and W.L. Staudenbauer **Plasmid R1 DNA replication dependent on protein synthesis in cell-free extracts of *E.coli*** 326

R.H. Lange **Highly conserved lipoprotein assembly in teleost and amphibian yolk-platelet crystals** 329

BOOK REVIEWS

Kenneth E. Boulding **The Lean Years: Politics in the Age of Scarcity (by Richard Barnet). Energy: from Opulence to Sufficiency (by Jerrold H. Krenz)** 331

Peter W. J. Rigby **DNA Tumor Viruses. Molecular Biology of Tumor Viruses, Part 2, 2nd Edn (edited by J. Tooze)** 332

Paul Davies **Gravity, Particles, and Astrophysics (by Paul S. Wesson)** 332

Kenneth Blaxter **Britain's Future in Farming (edited by Sir Frank Engledow and Leonard Amey)** 333

John Yarwood **A User's Guide to Vacuum Technology (by John F. O'Hanlon)** 333

Philip C. Hanawalt **Biological Effects of Ultraviolet Radiation (by Walter Harm)** 333

John Yudkin **Consuming Passions: The Anthropology of Eating (by Peter Farb and George Armelagos)** 334

MISCELLANY

100 years ago 226

Corrigenda and erratum 330

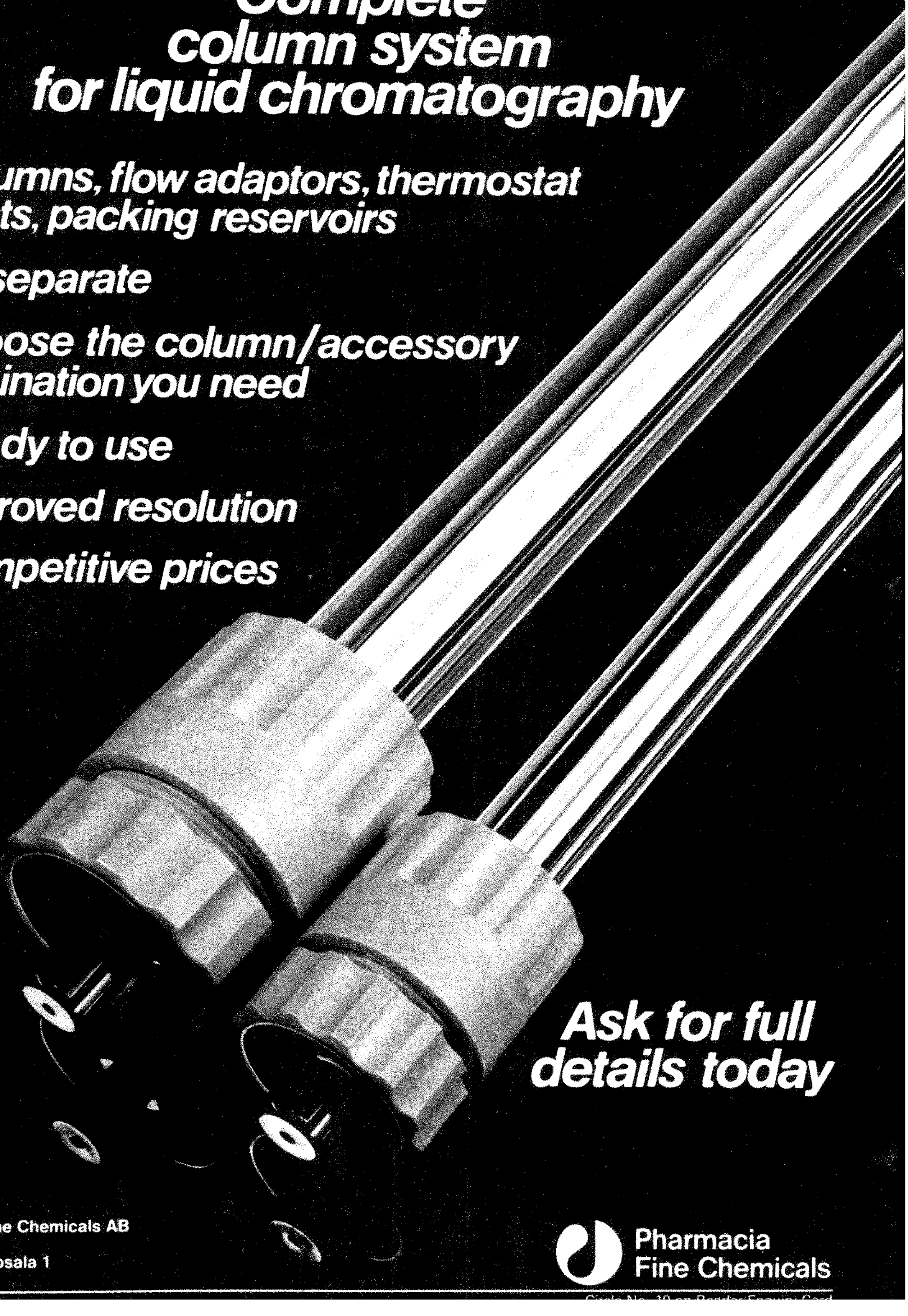
Announcements 336

Author Index Vol.288 ix

New on the Market xvii

Classified advertising xviii

***Nature* Indexes:** This week *Nature* includes an author index Volume 288, covering November and December 1980. Previous author indexes for 1980 appeared in the issues of 20 March (Vol. 283), 22 May (Vol.284), 24 July (Vol.285), 18 September (Vol.286) and 20 November (Vol.287).



NEW
Complete column system for liquid chromatography

- Columns, flow adaptors, thermostat jackets, packing reservoirs
- All separate
- Choose the column/accessory combination you need
- Ready to use
- Improved resolution
- Competitive prices

Ask for full details today

harmacia Fine Chemicals AB
ox 175
-751 04 Uppsala 1
weden

Pharmacia
Fine Chemicals

nature

Volume 289 No.5796 29 January 1981

NEWS

Madrid, Helsinki and exchanges 337
Naming names in British universities 337
Johns Hopkins wins new space institute: Princeton loses space telescope ground base 339
Dutch partners for La Palma observatory 339
Toxic waste: Dutch dumps 340
Genetic engineering: Planning bugs 340
Lamont director quits 341
Soviet fishing: Whither whaling? 341
Pneumoconiosis unit: Dusty answers 342
Bulgar birthday: Two for one 342
Reef aggregation 342
Brazilian ecology: Polluting by decree 342
Security conference: Back to Madrid 343
Indians home from home 343
Correspondence: Galileo's trial (H. Reboul); Don't be vague (D. Pike); Evolutionary thought (M. Howgate); Engels on Darwin (H. Crabtree); Motorbike safety (G. K. McGinty). 344

NEWS AND VIEWS

The Nobel awards for medicine or physiology: Sir Peter Medawar comments 345
Meteorite breakup: Michael R. Dence examines explanations for crater groupings 346
Flamingos, stilts and whales: Andrew Milner looks at flamingo systematics and feeding 346
Cement — a respectable material? P.D. Double thinks cement can get stronger 348
Understanding interatomic forces: Gillian Peach on the analysis of spectral lines 349
Current issues in conservation: Jared M. Diamond on problems of reserve management 350
Chlorinated dioxins and the environment: Alistair Hay reports a recent meeting 351

REVIEW ARTICLE

J. Cairns **The origin of human cancers** 353

ARTICLES

H.W. Nesbitt, G.M. Bancroft, W.S. Fyfe, S.N. Karkhanis, A. Nishijima and S. Shin **Thermodynamic stability and kinetics of perovskite dissolution** 358
W.H. Lehn and I. Schroeder **The Norse merman as an optical phenomenon** 362
I.A. Wilson, J.J. Skehel and D.C. Wiley **Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution** 366
D.C. Wiley, I.A. Wilson and J.J. Skehel **Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation** 373
G. Chu and P.A. Sharp **A gene chimaera of SV40 and mouse β-globin is transcribed and properly spliced** 378

LETTERS

R. Woo and A. Ishimaru **Eddy diffusion coefficient for the atmosphere of Venus from radio scintillation measurements** 383
D. Wills and B.J. Wills **Optical observations of OV–236** 384
S.S. Prasad **Excited ozone is a possible source of atmospheric N_2O** 386
J.D. Birchall, A.J. Howard and K. Kendall **Flexural strength and porosity of cements** 388
J.K. Jeszka, J. Ulański and M.Kryszewski **Conductive polymer: reticulate doping with charge-transfer complex** 390
P. Cloos, C. Badot and A. Herbillon **Interlayer formation of humin in smectites** 391

Contents continued overleaf

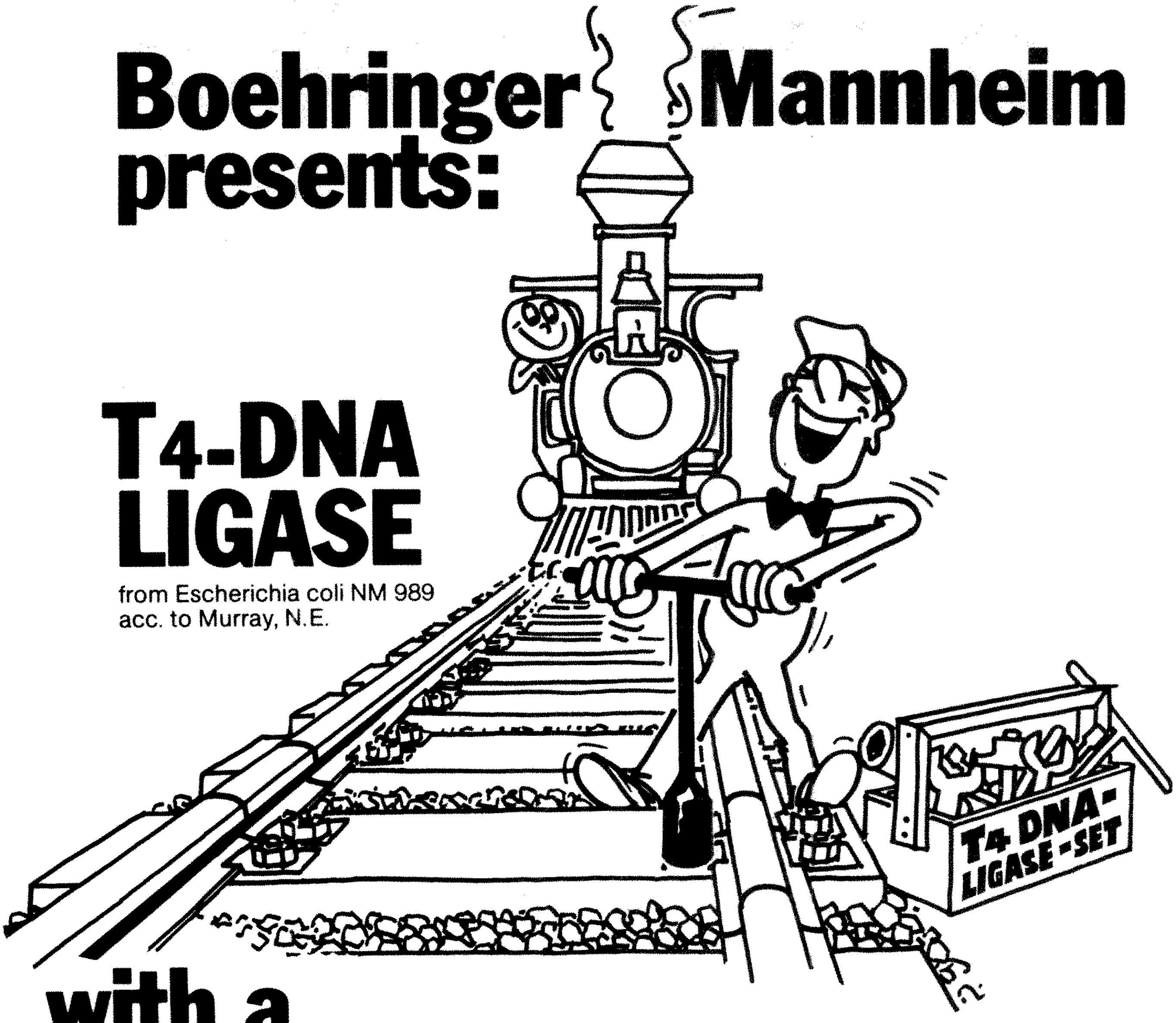
Cover: The complete structure of the influenza haemagglutinin molecule and stages in its construction. The amino acid changes associated with new epidemic strains of influenza can be mapped onto four major sites on the haemagglutinin structure (see pages 366, 373 and 422).

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

Boehringer Mannheim presents:

T4-DNA LIGASE

from Escherichia coli NM 989
acc. to Murray, N.E.



with a

Specific activity: 3000 units*/mg Protein

*acc. to Weiss B. et al (1968) J. Biol. Chem. **243**, 4543–4555

checked for the absence of
endonuclease
5'-exonuclease and 5'-phosphatase
3'-exonuclease

is now available under the Cat. No. 481 220

Please ask your local representative
for the DNA-Ligase brochure and for detailed information
or

mannheim boehringer

Boehringer Mannheim GmbH
Biochemica
P.O. Box 310120
D-6800 Mannheim
W. Germany

Circle No.07 on Reader Enquiry Card.

LETTERS

D.R. Fearn and D.E. Loper	**Compositional convection and stratification of Earth's core**	393
N.B.W. Harris and I.G Gass	**Significance of contrasting magmatism in North East Africa and Saudi Arabia**	394
R. Bhardwaj, R L. Pan and E.L. Gross	**Solar energy conversion by chloroplast photoelectrochemical cells**	396
I.J. Miller and S.K. Fellows	**Liquefaction of biomass as a source of fuels or chemicals**	398
M.P. Cochrane and C.M. Duffus	**Endosperm cell number in barley**	399
M.N. Perkins, N.J. Rothwell, M.J. Stock and T.W. Stone	**Activation of brown adipose tissue thermogenesis by the ventromedial hypothalamus**	401
E.P. Wallen and F.W. Turek	**Photoperiodicity in the male albino laboratory rat**	402
P.N. Goodfellow and P. Tippett	**A human quantitative polymorphism related to Xg blood groups**	404
D.J.Schrier, J.L. Sternick, E.M. Allen and V.L. Moore	**Genetic basis of BCG-induced suppression of delayed hypersensitivity**	405
J.C. Lee and J.N. Ihle	**Chronic immune stimulation is required for Moloney leukaemia virus-induced lymphomas**	407
D.R. Shaw and F.M. Griffin Jr	**Phagocytosis requires repeated triggering of macrophage phagocytic receptors during particle ingestion**	409
R.W. Kullberg, P. Brehm and J.H. Steinbach	**Nonjunctional acetylcholine receptor channel open time decreases during development of *Xenopus* muscle**	411
D.H.Wreschner, J.W. McCauley, J.J. Skehel and I.M. Kerr	**Interferon action — sequence specificity of the ppp(A2′p)$_n$A-dependent ribonuclease**	414
R. Sklar and B. Strauss	**Removal of O^6-methylguanine from DNA of normal and xeroderma pigmentosum-derived lymphoblastoid lines**	417
E. Ungewickell and D. Branton	**Assembly units of clathrin coats**	420
M. Waterfield, G. Scrace and J. Skehel	**Disulphide bonds of haemagglutinin of Asian influenza virus**	422

BOOK REVIEWS

Eric Ashby	**The Politics of Self-Sufficiency (by Michael Allaby and Peter Bunyard)**	425
T.A. Connors	**Ethylene Dichloride: A Potential Health Risk? 5th Banbury Report (edited by B. Ames, P. Infante and R. Reitz)**	426
Colin Upstill	**An Introduction to Catastrophe Theory (by P.T. Saunders)**	426
J.M. Irvine	**Nuclear Spectroscopy (edited by G. F. Bertsch and D. Kurath)**	427
G.D. Heathcote	**Vectors of Plant Pathogens (edited by Kerry F. Harris and Karl Maramorosch)**	427
J.B. Nelson	**Behaviour of Marine Animals. Vol. 4 Marine Birds (edited by J. Burger, B.L. Olla and H.E. Winn)**	427
Howard R. Morris	**Biochemical Applications of Mass Spectrometry: First Supplementary Volume (edited by George R. Waller and Otis C. Dermer)**	428

MISCELLANY

Books received 429

Classified advertising xix

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership Many reviews are invited, but submitted articles may also be accepted, it is advisable to consult us before writing a review article
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables), they are reports of major research developments
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible

Manuscripts may be submitted either to London or New York Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced The title should be brief and informative Pages should be numbered References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends Units should be identified in the margin on their first appearance Equations should occupy single lines if possible exp (x) is preferred to e^x if 'x' is more than one character

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials Most figures are reduced to one column width so originals should be about as wide as a page of *Nature* To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted We are always glad to see artwork for possible use on the cover, but cannot guarantee its return

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

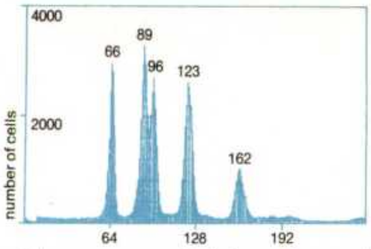
Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov (1976)

The Ortho ICP 22A Flow Cytometer utilizes a high power mercury arc lamp in combination with a Kohler optical system to achieve extraordinary precision and wide versatility. The capabilities of this instrument are best evaluated from the standpoint of the results it can provide. Examine the actual fluorescence histograms presented.

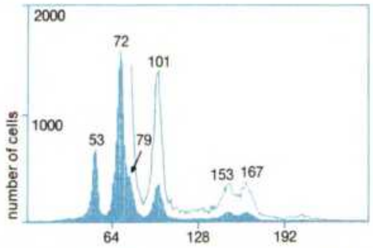
The high resolution of the instrument is complemented by its dual channel fluorescence detectors, which allow correlated measurements of two different dyes. Thus, the ICP 22A can provide simultaneous measurement of per cell DNA and protein content. Because of its Kohler illumination system it can even measure the DNA content of cells in which the DNA is irregularly distributed. Given the range of specific stoichiometric dyes available, its versatility is almost unmatched. To enhance that versatility still further, the ICP 22A will soon offer electronic cell volume measurement in addition to the other parameters (available September 1980).

Therapy monitoring by flow cytometric analysis

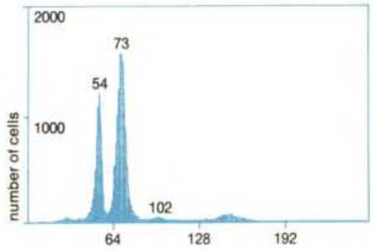
1. DNA histograms show a multiclonal tumor with four distinct tumor cell lines. Left peak in all histograms is normal cells. During three month chemotherapy with bleomycin and DTIC, the three higher ploidy tumor lines disappeared leaving only the more resistant lowest ploidy line. This biopsy series shows the different sensitivities of tumor cell lines in the same patient.



1a. At commencement of chemotherapy, the DNA histogram shows four cell lines with different ploidy stages. Left peak is normal diploid cells.



1b. After two months the DNA histogram shows suppression of three of the tumor cell lines.



1c. One month later, the DNA histogram shows the disappearance of higher ploidy tumor cell lines and the resistance of lower ploidy tumor cell line to chemotherapy.

In order to provide customers with complete support, Ortho maintains an Applications Laboratory that is available for confirmation and consultation services. Additionally, an international network of service technicians is ready to aid customers should calibration or repairs ever be required.

The Ortho ICP 22A offers remarkable precision and versatility at an exceptionally affordable price. We invite you to contact the Research Instruments Division of Ortho to find out more about the ICP Flow Cytometer. It's available either as a stand alone flow cytometric analyzer or as part of an Ortho System 40 or 60, with dual laser illumination.

Protocols No. 28 and 29

Two new Protocols; *One step Ethidium Bromide DNA/Cell Staining Procedure* and *DNA analysis using D.A.P.I.* are available from Ortho. Write or call for your complimentary copies. (617) 329-6100

ORTHO INSTRUMENTS
410 UNIVERSITY AVE., WESTWOOD, MA 02090

Precision, versatility and exceptional value.

The ICP Flow Cytometer

Circle No. 02 on Reader Enquiry Card.

nature

Volume 289 No.5797 5 February 1981

NEWS

Two views of the causes of cancers 431

University empire confronts a crisis 432

Cutback on West German development: Money now in short supply in Bonn 433

Polish students sit in at universities 433

Telecommunications suit: Pact postponed 434

European Community: Squeezing grants 434

UN energy conference: Prospects improve 434

Ariane space launcher: Still in trouble 435

UK research councils: Getting off lightly 435

Canadian research: Aiming high 436

Environmental cadmium: Europe to ban? 436

Soviet psychology: IQ rehabilitated? 437

Orlov hunger strike 437

Correspondence: Laser war games (E. Wheeler); Exhibition policy (V. Demoulin, R. L. Gregory); Corrigendum (F. H. Malpress); Soviet rumbles (A. A. Meyer) 438

NEWS AND VIEWS

One gene's intron is another gene's exon: Piet Borst and Leslie A. Grivell on a mitochondrial RNA processing protein encoded by an intron 439

Beetles in ballast: R.A. Crowson on involuntary colonists 441

A bright nearby supernova: Roger A. Chavalier describes observations of supernova 1979c 441

Negative cooperativity at the insulin receptor: Alexander Levitzki examines the evidence 442

The Trigger experiment: Daniel W. Swift and Hans C. Stenbaeck-Nielsen describe attempts to simulate auroral processes 443

Are membrane proteins introverted?: N. Michael Green describes an 'inside-out' protein 444

African lithosphere: Paul Mohr on the thickness of the African plate 445

Fibroblast interferon: D.A.J. Tyrell on an anti-interferon globulin in man 446

ARTICLES

A. Brahic and B. Sicardy **Apparent thickness of Saturn's rings** 447

M. Matson and D.R. Wiesnet **New data base for climate studies** 451

W.R.A. Brown, A.N. Barclay, C.A. Sunderland and A.F. Williams **Identification of a glycophorin-like molecule at the cell surface of rat thymocytes** 456

L.J. Korn and J.B. Gurdon **The reactivation of developmentally inert 5S genes in somatic nuclei injected into *Xenopus* oocytes** 461

N. Panayotatos and R.D. Wells **Cruciform structures in supercoiled DNA** 466

LETTERS

C.G.T. Haslam, S. Kearsey, J.L. Osborne, S. Phillipps and H. Stoffel **Galactic γ and radio synchrotron radiation** 470

L. Thomas, C.P. Chaloner and S.K. Bhattacharyya **Laser–radar measurements in southern England of aerosols from Mount St Helens** 473

P.H. Gaskell **Similarities in amorphous and crystalline transition metal–metalloid alloy structures** 474

L. Petrakis and D.W. Grandy **Formation and behaviour of coal free radicals in pyrolysis and liquefaction conditions** 476

K.P. Games and M.E. Baker **Determination of geomagnetic archaeomagnitudes from clay pipes** 478

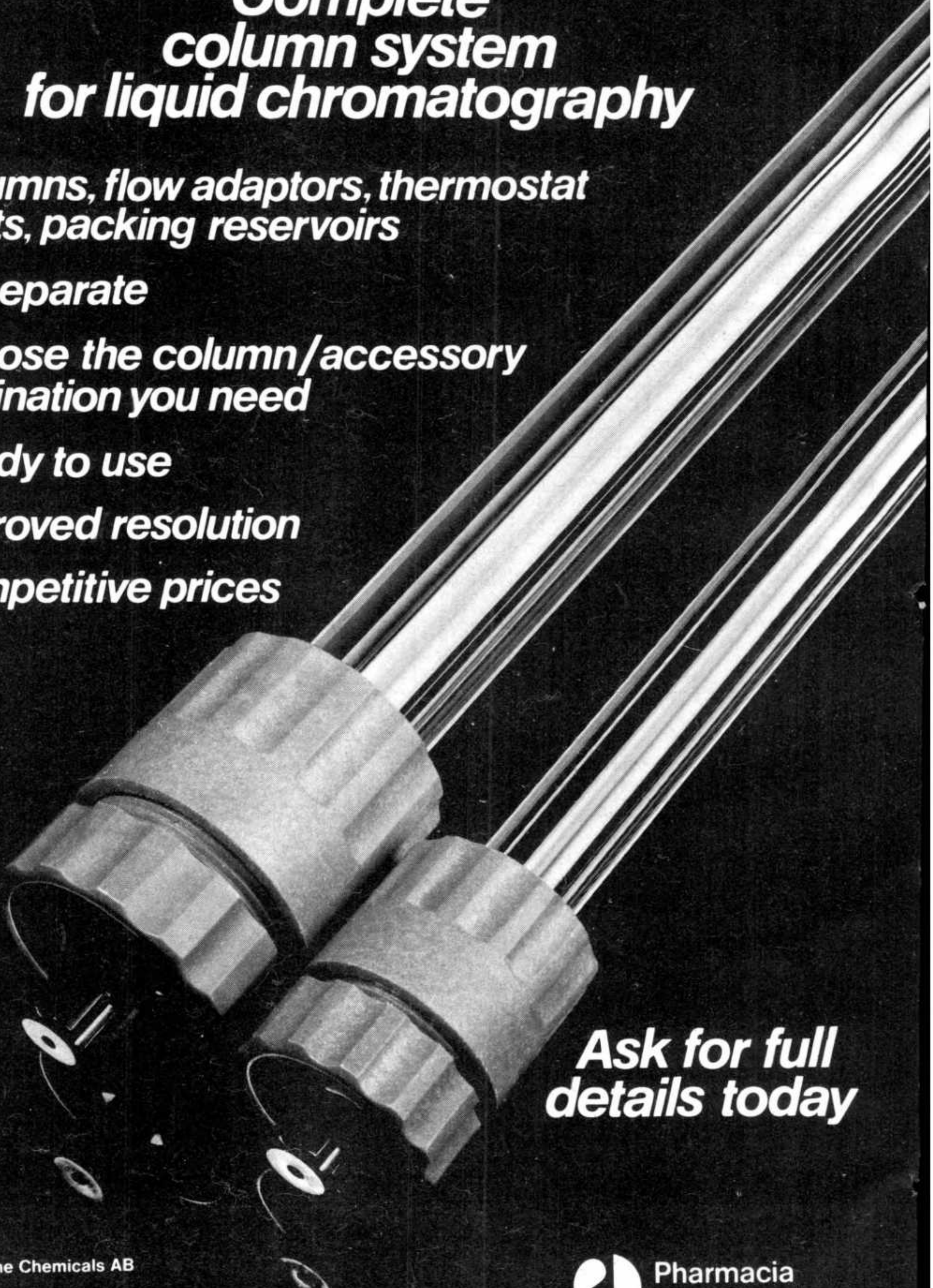
A.G. Wintle **Thermoluminescence dating of late Devensian loesses in southern England** 479

Contents continued overleaf

Cover: A computer reconstruction of mean snow cover extent of the Northern Hemisphere for December 1979 prepared by Marilin Varnadore of NOAA's National Earth Satellite Service. Snow cover is depicted in gold. The reconstruction is based on satellite data discussed by Matson and Wiesnet on page 451.

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/06000C-02$01.00



NEW
Complete column system for liquid chromatography

- *Columns, flow adaptors, thermostat jackets, packing reservoirs*
- *All separate*
- *Choose the column/accessory combination you need*
- *Ready to use*
- *Improved resolution*
- *Competitive prices*

Ask for full details today

Pharmacia Fine Chemicals AB
Box 175
S-751 04 Uppsala 1
Sweden

Pharmacia
Fine Chemicals

Circle No. 10 on Reader Enquiry Card.

R.S. Harmon, L.S. Land, R.M. Mitterer, P. Garrett, H.P. Schwarcz and G.J. Larson — **Bermuda sea level during the last interglacial** 481

G.H. Rau — **Low $^{15}N/^{14}N$ in hydrothermal vent animals: ecological implications** 484

N. Healy-Williams and D.F. Williams — **Fourier analysis of test shape of planktonic foraminifera** 485

T.H. Clutton-Brock, S.D. Albon and F.E. Guinness — **C arental investment in male and female offspring in polygynous mammals** 487

J.F. Ashmore and G. Falk — **Photon-like signals following weak rhodopsin bleaches** 489

M.M.B. Kay — **Isolation of the phagocytosis-inducing IgG-binding antigen on senescent somatic cells** 491

I. Ando and M. Hurme — **Self-MHC-restricted cytotoxic T-cell response without thymic influence** 494

A. Vallbracht, J. Treuner, B. Flehmig, K.-E. Joester and D. Niethammer — **Interferon-neutralizing antibodies in a patient treated with human fibroblast interferon** 496

A. Björklund, U. Stenevi, S.B. Dunnett and S.D. Iversen — **Functional reactivation of the deafferented neostriatum by nigral transplants** 497

M. Sar and W.E. Stumpf — **Central noradrenergic neurones concentrate ^{3}H-oestradiol** 500

R. Martin and K.H. Voigt — **Enkephalins co-exist with oxytocin and vasopressin in nerve terminals of rat neurohypophysis** 502

D.W. Lincoln and M.B. Renfree — **Milk ejection in a marsupial, *Macropus agilis*** 504

J.B. Johnson and K.S. Hagen — **A neuropterous larva uses an allomone to attack termites** 506

G. Aguilera, A. Schirar, A. Baukal and K.J. Catt — **Circulating angiotensin II and adrenal receptors after nephrectomy** 507

P.R. Stanfield, F.M. Ashcroft and T.D. Plant — **Gating of a muscle K^+ channel and its dependence on the permeating ion species** 509

D. Shields, T.G. Warren, S.E. Roth and M.J. Brenner — **Cell-free synthesis and processing of multiple precursors to glucagon** 511

A.J. Moody, J.J. Holst, L. Thim and S. Lindkaer Jensen — **Relationship of glicentin to proglucagon and glucagon in the porcine pancreas** 514

D. Sullivan, N. Brisson, B. Goodchild, D.P.S. Verma and D.Y. Thomas — **Molecular cloning and organization of two leghaemoglobin genomic sequences of soybean** 516

BOOK REVIEWS

Alvin M. Weinberg — **The Fast Breeder Reactor: Need? Cost? Risk? (edited by Colin Sweet)** 519

Brian Bertram — **Conservation Biology: An Evolutionary–Ecological Perspective (edited by Michael E. Soulé and Bruce A. Wilcox)** 520

William R. Cotton and Kevin R. Knupp — **Clouds and Storms: The Behaviour and Effect of Water in the Atmosphere (by F. H. Ludlam)** 520

H.A.O. Hill — **An Introduction to Spectroscopy for Biochemists (edited by S. B. Brown)** 521

J. Hodgkin — **Nematodes as Biological Models, Vols 1 and 2 (edited by Bert M. Zuckerman)** 521

Donald A. Levin — **Demography and Evolution in Plant Populations (edited by Otto T. Solbrig)** 522

B.A. Pethica — **The Hydrophobic Effect: Formation of Micelles and Biological Membranes, 2nd Edn (by Charles Tanford)** 522

MISCELLANY

New on the market xvi

Classified advertising xx

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: exp (a) is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

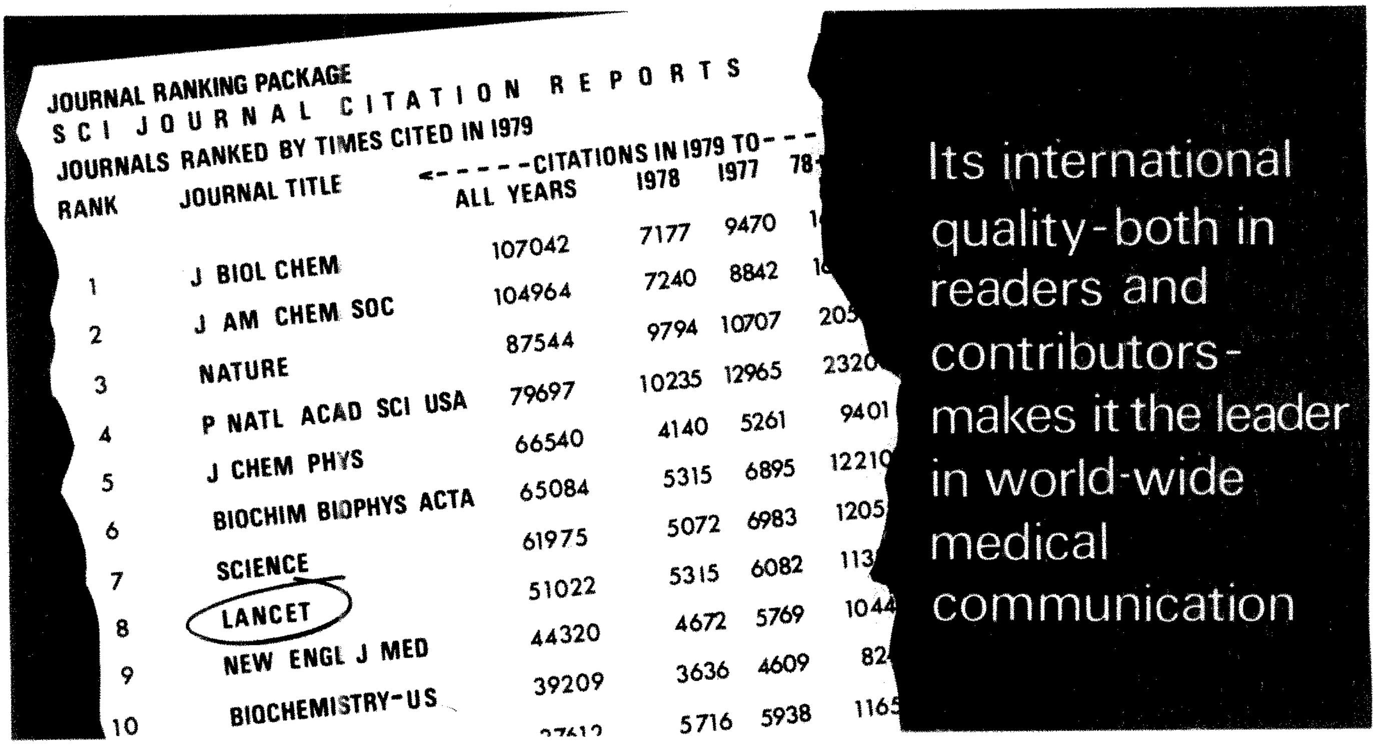
In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov. (1976).

THE LANCET

On the frontiers of medical and scientific knowledge THE LANCET is not only the last word - it's the first!

Consistently ranked by the Institute of Scientific Information as the most cited medical journal in the world.



JOURNAL RANKING PACKAGE
SCI JOURNAL CITATION REPORTS
JOURNALS RANKED BY TIMES CITED IN 1979
<-----CITATIONS IN 1979 TO---

RANK	JOURNAL TITLE	ALL YEARS	1978	1977	78
1	J BIOL CHEM	107042	7177	9470	1
2	J AM CHEM SOC	104964	7240	8842	1
3	NATURE	87544	9794	10707	205
4	P NATL ACAD SCI USA	79697	10235	12965	2320
5	J CHEM PHYS	66540	4140	5261	9401
6	BIOCHIM BIOPHYS ACTA	65084	5315	6895	12210
7	SCIENCE	61975	5072	6983	1205
8	LANCET	51022	5315	6082	113
9	NEW ENGL J MED	44320	4672	5769	1044
10	BIOCHEMISTRY-US	39209	3636	4609	82
		[illegible]	5716	5938	1165

Its international quality - both in readers and contributors - makes it the leader in world-wide medical communication

ORDER FORM

To – THE LANCET LTD., 7 Adam Street, Adelphi, London WC2N 6AD.
Please send me THE LANCET each week for one year.

☐ £27.00 in UK and EIRE ☐ Cheque herewith
☐ £40.00 Overseas (Airspeeded) ☐ Invoice me
☐ £12.00 special rate*, (Available in UK and Eire only) for doctors within 5 years of registration or graduation and for medical students.

*Date of registration or student at – ____________________

BLOCK CAPITALS PLEASE

NAME ____________________

ADDRESS ____________________

SIGNATURE ____________________ DATE ____________________

nature

Volume 289 No.5798 12 February 1981

NEWS

Who now believes in the neutron bomb? 523

New Reagan budget a mixed blessing: Social science and planetary missions cut 524

European high-energy physics cuts 524

Soviet agriculture: Soil first 525

Solar power: Crimean costs 525

Wildlife conservation: Keepers' charter 525

German nuclear power: Local difficulties 526

EEC nuclear research: Reactor guarded 526

Data protection laws: UK lags behind 527

European astronomy: Winning ways 527

Correspondence: Selfish genes in race or politics (R. Dawkins) 528

NEWS AND VIEWS

A new measure of intelligence?: N.J. Mackintosh on IQ and inspection time 529

Meteor train luminosity: W.J. Baggaley explains the significance of spectroscopic observations 530

Selection and induction in the development of autonomic neurones: Miranda Robertson reports a recent conference 532

Quasar alignments: chance or necessity: M.G. Edmunds looks at some extraordinary quasar configurations 533

Antimatter back to front: C.H. Llewellyn Smith on the history of CP violation 534

Genetic engineering and foot and mouth disease vaccines: J.B. Brooksby comments 535

Well-groomed predecessors: R.D. Martin explains the significance of the toothcomb of lemurs and lorises 536

REVIEW ARTICLE

A. Cheramy, V. Leviel and J. Glowinski — **Dendritic release of dopamine in the substantia nigra** 537

ARTICLES

R.J. Prance, A.P. Long, T.D. Clark, A. Widom, J.E. Mutton, J. Sacco, M.W. Potts, G. Megaloudis and F. Goodall — **Macroscopic quantum electrodynamic effects in a superconducting ring containing a Josephson weak link** 543

B. Zipser and R. McKay — **Monoclonal antibodies distinguish identifiable neurones in the leech** 549

H. Küpper, W. Keller, C. Kurz, S. Forss, H. Schaller, R. Franze, K. Strohmaier, O. Marquardt, V.G. Zaslavsky and P.H. Hofschneider — **Cloning of cDNA of major antigen of foot and mouth disease virus and expression in *E. coli*** 555

K. Nakano, I. Nakamura and G. Cudkowicz — **Generation of F_1 hybrid cytotoxic T lymphocytes specific for self H-2** 559

R. Risser and D.J. Grunwald — **Production of anti-self H-2 antibodies by hybrid mice immune to a viral tumour** 563

LETTERS

S. Danaher, D.J. Fegan, N. A. Porter and T. C. Weekes — **γ-Ray observations of Cygnus X-3 at energies of 10^{12} eV** 568

M. Lugg — **Effect of a surrounding star cluster on a magnetoid** 569

J.M. Prospero, R.A. Glaccum and R.T. Nees — **Atmospheric transport of soil dust from Africa to South America** 570

Y. Yamamoto, N. Baba and S. Tajima — **Coloured materials and photoluminescence centres in anodic film on aluminium** 572

Contents continued overleaf

Cover: A leech ganglion stained with the monoclonal antibody Lan3-2 and rhodamine-labelled second antibody. The fluorescence micrograph shows the nociceptive cell bodies and processes stained. Zipser and McKay report on page 549 the use of monoclonal antibodies to distinguish identifiable neurones as well as numerous unidentified cells in the leech nervous system.

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/07000C-02$01.00

Raven Press

New Titles in Neurosciences

The Primary Acoustic Nuclei

Rafael Lorente de Nó

Written by one of the world's most outstanding neuroanatomists, this monograph contains reproductions of nearly 100 original drawings prepared from Golgi or Golgi-Cox sections, as well as microphotographs of 19 myelin-stained sections that present the locations of the various subdivisions of the acoustic nuclei and the courses of the main association tracts. **200 pp., 8½ x 11 inches, 129 figures, $25.00 (1981) ISBN 0-89004-318-3**

Kuru: Early Letters and Field-Notes from the Collection of D. Carleton Gajdusek

Editors: Judith Farquhar and D. Carleton Gajdusek

This volume is a collection of D. Carleton Gajdusek's writings during the first year of his medical investigation of Kuru. The investigation of this fatal disease provides illuminating insights into the scientific method and the history and philosophy of science, as well as valuable leads for future work on slow virus infections and chronic and heredofamilial degenerative diseases of man's central nervous system.
366 pp., $32.00 (1981) ISBN 0-89004-359-0

Amino Acid Neurotransmitters

Editors: Francis V. DeFeudis and Paul Mandel

This volume places particular emphasis on roles of amino acid in integrated central nervous system functions and behavior, receptor pharmacology at both *in vivo* (electrophysiological) and *in vitro* (ligand-binding) levels, amino acid uptake and release, enzymology metabolism, and autoradiographic studies.
600 pp., $54.00 (1981) ISBN 0-89004-595-X
ISBN 0-89004-595-X

Psychobiology of Aggression and Violence

Editor: Luigi Valzelli

This monograph by an internationally renowned neuroscientist integrates findings from animal and human studies to provide insights into the physiology and pathology of aggression, and to formulate potential controls for individuals and groups.
265 pp., $24.00 (1981) ISBN 0-89004-403-1

GABA and Benzodiazepine Receptors

Advances in Biochemical Psychopharmacology, Vol. 26

Editors: E. Costa, Gaetano Di Chiara, and G. L. Gessa

This book examines GABA participation in the action of benzodiazepines: the structure of the GABA receptor and the benzodiazepine binding site, the localization of these receptors within specific brain regions, structure-activity relationships with agonists, and the involvement of GABA receptors in specific functional pathways.
321 pp., $36.00 (1981) ISBN 0-89004-530-5

Glutamate as a Neurotransmitter

Advances in Biochemical Psychopharmacology, Vol. 27

Editors: G. Di Chiara and G. L. Gessa

Basic neurobiological evidence, as well as the pharmacological and clinical implications of the role of glutamate as transmitter are broadly and systematically treated, from the anatomy, biochemistry, and electrophysiology of glutamatergic synapses and receptors to the role of glutamate analogs, particularly kainic acid, in neuropathology. **464 pp., $41.00 (1981) ISBN 0-89004-420-1**

Neurosecretion and Brain Peptides: Implications for Brain Function and Neurological Disease

Advances in Biochemical Psychopharmacology, Vol. 28

Editors: Joseph B. Martin, Seymour Reichlin, and Katherine L. Bick

This volume presents the most recent findings regarding the synthesis, anatomic distribution, and function of neuropeptides and provides a critical evaluation of their importance for the understanding of the pathogenesis of neurologic disease.
725 pp., $58.00 (1981) ISBN 0-89004-535-6

Journal of Cerebral Blood Flow and Metabolism

Quarterly, Vol.1 (1981) ISSN-0271-678X
Subscription rates: Personal subscriptions-$85.00, U.S.A.;$95.00 elsewhere.
Library and institutional subscription-$95.00, U.S.A.; $105.00 elsewhere.
This quarterly journal is the official publication of the newly formed International Society of Cerebral Blood Flow and Metabolism. This publication offers new information on experimental, theoretical, and clinical aspects of brain circulation and metabolism.

Circle No.02 on Reader Enquiry Card.

Order Form

Please send: ______

Prices subject to change without notice. Prices slightly higher outside the U.S. and Canada.

☐ To save postage and handling charges, I enclose payment (or charge my credit card below). N.Y. residents add 8% sales tax.
Check one: ☐ American Express ☐ VISA ☐ Diners Club
☐ Master Charge (4-digit interbank no.: ☐☐☐☐)
Credit card number: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐
Card expiration date: ______
Name (please print) ______
Address ______
Country ______ Postal Code ______

Raven Press **1140 Avenue of the Americas, New York, N.Y. 10036 U.S.A.**

Na 21281.

N. Shimizu **Trace element incorporation into growing augite phenocryst** 575

D.T. Aldiss **Plagiogranites from the ocean crust and ophiolites** 577

J.R.L. Allen **Lower Cretaceous tides revealed by cross-bedding with mud drapes** 579

S.R. Aston and D.A. Stanners **Plutonium transport to and deposition and immobility in Irish Sea intertidal sediments** 581

K.D. Rose, A. Walker and L.L. Jacobs **Function of the mandibular tooth comb in living and extinct mammals** 583

L.L. Jacobs **Miocene lorisid primates from the Pakistan Siwaliks** 585

D.W.M. Leung and J.D. Bewley **Immediate phytochrome action in inducing α-galactosidase in lettuce seeds** 587

J.G. Lopez and G.R. Vela **True morphology of the Azotobacteraceae — filterable bacteria** 588

A. Imada, K. Kitano, K. Kintaka, M. Muroi and M. Asai **Sulfazecin and isosulfazecin, novel β-lactam antibiotics of bacterial origin** 590

E.R. Tovey, M.D. Chapman and T.A.E. Platts-Mills **Mite faeces are a major source of house dust allergens** 592

F. Audibert, M. Jolivet, L. Chedid, J.E. Alouf, P. Boquet, P. Rivaille and O. Siffert **Active antitoxic immunization by a diphtheria toxin synthetic oligopeptide** 593

H.A. Drexhage, G.F. Bottazzo, L. Bitensky, J. Chayen and D. Doniach **Thyroid growth-blocking antibodies in primary myxoedema** 594

D.C. Ogden, S.A. Siegelbaum and D. Colquhoun **Block of acetylcholine-activated ion channels by an uncharged local anaesthetic** 596

A. Saito, J.A. Williams and I.D. Goldfine **Alterations in brain cholecystokinin receptors after fasting** 599

P.W. Wilson and D.E.M. Lawson **Vitamin D-dependent phosphorylation of an intestinal protein** 600

J.S. Edmonds and K.A. Francesconi **Arseno-sugars from brown kelp *(Ecklonia radiata)* as intermediates in cycling of arsenic in a marine ecosystem** 602

B.W. Dijkstra, J. Drenth and K.H. Kalk **Active site and catalytic mechanism of phospholipase A_2** 604

R. Wetzel **Assignment of the disulphide bonds of leukocyte interferon** 606

B.-Z. Shilo and R.A. Weinberg **Unique transforming gene in carcinogen-transformed mouse cells** 607

K. Nyman, K. Nakamura, H. Ohtsubo and E. Ohtsubo **Distribution of the insertion sequence IS1 in Gram-negative bacteria** 609

BOOK REVIEWS

John C. Marshall **Language and Learning: The Debate Between Jean Piaget and Noam Chomsky (edited by M. Piattelli-Palmarini)** 613

D.V. Bugg **Theory of Meson Interactions with Nuclei (by Judah M. Eisenberg and Daniel S. Koltun)** 614

Stanwyn G. Shetler **The Arctic and Antarctic: Their Division into Geobotanical Areas (by V.D. Aleksandrova)** 615

M.J. Seaton **Atomic Structure (by E.U. Condon and Halis Odabaşi)** 615

Toshisada Nishida **Animal Tool Behavior: The Use and Manufacture of Tools by Animals (by Benjamin B. Beck)** 616

MISCELLANY

Announcements 617

Classified advertising xvi

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: exp (a) is preferred to e^{a} if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov. (1976).

NICK TRANSLATION KIT

5′ NICK$_{OH}$ TRANSLATION KIT 3′

3′ 5′

Now from Amersham

Tritium Nick Translation Kit TRK.700

A new kit for the preparation of DNA labelled with tritium to the specific activity of your choice.
Recommended protocols give 5×10^6–4×10^7 dpm/μg DNA
The kit comprises a set of non-radioactive reagents + 1 mCi [1′, 2′, 5-^{3}H]dCTP.

Already available from stock

PB.5100 non-radioactive reagents + 1 mCi [α-^{32}P]dCTP
PB.5025 non-radioactive reagents + 250 μCi [α-^{32}P]dCTP
for the preparation of DNA labelled with ^{32}P at up to 4×10^8 dpm/μg DNA

N.5000 non-radioactive reagents only

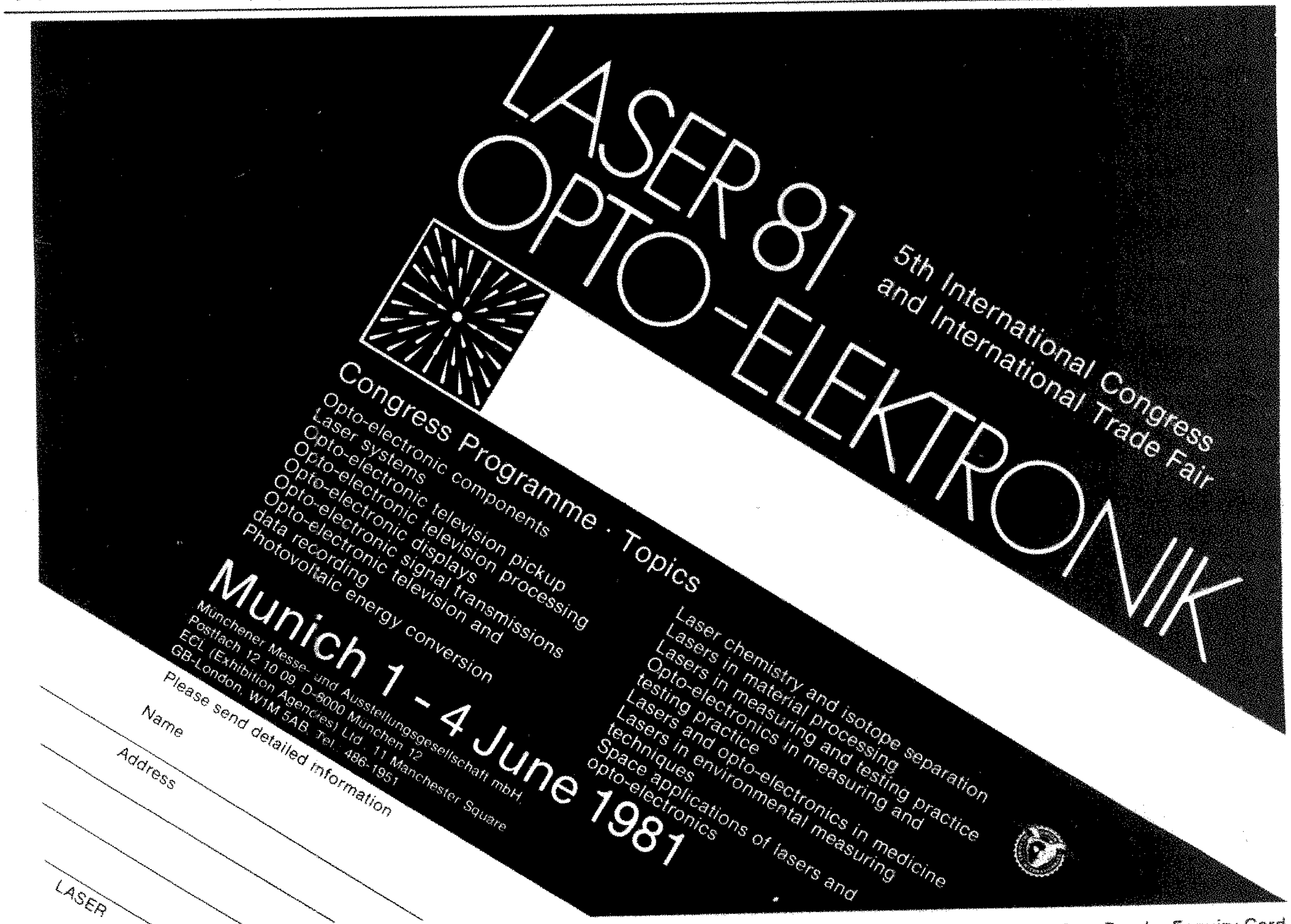
for the preparation of DNA labelled with the nucleotide of your choice.

Full details of the kit and protocols are available on request.

The Radiochemical Centre
Amersham

The Radiochemical Centre Limited, Amersham, England. Telephone: 024-04-4444.
In the U.S.A. and Canada: Amersham Corporation, Illinois 60005. Telephone: 312-364-7100
and 800-323-9750 (Tollfree).
In W. Germany: Amersham Buchler GmbH & Co KG, Braunschweig. Telephone: 05307-4691.

Circle No.04 on Reader Enquiry Card.



LASER 81
OPTO-ELEKTRONIK

5th International Congress
and International Trade Fair

Congress Programme · Topics

Opto-electronic components
Laser systems
Opto-electronic television pickup
Opto-electronic television processing
Opto-electronic displays
Opto-electronic signal transmissions
Opto-electronic television and data recording
Photovoltaic energy conversion

Laser chemistry and isotope separation
Lasers in material processing
Lasers in measuring and testing practice
Opto-electronics in measuring and testing practice
Lasers and opto-electronics in medicine
Lasers in environmental measuring techniques
Space applications of lasers and opto-electronics

Munich 1 - 4 June 1981

Münchener Messe- und Ausstellungsgesellschaft mbH,
Postfach 12 10 09, D-8000 München 12
ECL (Exhibition Agencies) Ltd., 11 Manchester Square
GB-London, W1M 5AB, Tel.: 486-1951

Please send detailed information

Name

Address

LASER

Circle No. 14 on Reader Enquiry Card.

nature

Volume 289 No.5799 19 February 1981

NEWS

Employment, training and opportunity 619
How to make dreams untrue 620
More secrecy on cryptography research: US academics lean towards self-restraint 621
Committee douches nuclear energy 621
European Community: Research compared 622
Toxic chemicals: UK regulations 622
Industrial innovation: States chip in 623
Telecommunications: Monopoly intact 623
Carcinogen criteria: US retracts 624
Greenhouse effect: Act now, not later 624
Plasma research: German setback 625
Tokamaks and stellarators 625
Princeton perplexities 625
Correspondence: Wrong rodents (M.F.W. Festing); French museums (P. Janvier); More museums (M.C. McKenna; G. Nelson); Genes and racism (E.O. Wilson); Origin of cancer (R.C. Garner and P.J. Hertzog) 626
The great badger debate: Lord Zuckerman 628

NEWS AND VIEWS

Too soon for the rehabilitation of Lamarck: a discussion of the results of Gorcynski and Steele (see on page 678) 631
Gamma-ray astronomy: R.D. Willis reports a recent conference 632
Diffusionless reactions and crystal engineering: J.M. Thomas on chemical reactions in the solid state 633
Habitat and rabies: P.J. Bacon and D.W. Macdonald on attempts to predict the spread of the disease 634
The expanding science of photobiology: John Jagger reports an international congress 636
Antisickling drugs and water structure: Ruth and Reinhold Benesch propose a new model of antisickling drug action 637
Fifty years of Slough's ionosphere: Henry Rishbeth celebrates the golden jubilee of ionospheric sounding 638

ARTICLES

E. Vincent, J.S. Killingley and W.H. Berger — **Stable isotope composition of benthic foraminifera from the equatorial Pacific** 639
O. Hagenbüchle, M. Tosi, U. Schibler, R. Bovey, P.K. Wellauer and R.A. Young — **Mouse liver and salivary gland α-amylase mRNAs differ only in 5′ non-translated sequences** 643
C.V. Wedmore and T.J. Williams — **Control of vascular permeability by polymorphonuclear leukocytes in inflammation** 646
P. Valenzuela, M. Quiroga, J. Zaldivar, W.J. Rutter, M.W. Kirschner and D.W. Cleveland — **Nucleotide and corresponding amino acid sequences encoded by α and β tubulin mRNAs** 650

LETTERS

E. Herbst — **Can negative molecular ions be detected in dense interstellar clouds?** 656
S.L. Shapiro and I. Wasserman — **Massive neutrinos, helium production and the primordial magnetic field** 657
M.E. Bailey and J. MacDonald — **Density evolution of radio sources fuelled by stellar mass loss** 659
R.H. Lacombe and B.M.J. Kellner — **Thermal-mechanical properties of a monomolecular film detected by a vibrating reed** 661
E.D. Hondros and C. Lea — **Grain boundary microchemistry and stress-corrosion failure of mild steel** 663
R.H. Sibson — **Controls on low-stress hydro-fracture dilatancy in thrust, wrench and normal fault terrains** 665
J. Thiede — **Reworking in Upper Mesozoic and Cenozoic central Pacific deep-sea sediments** 667

Cover: An Augergraph depicting the microchemistry of a grain boundary at the crack front of a steel failing by stress corrosion. The analysis is made by Auger electron spectroscopy which monitors a depth of only one or two atom layers. The yellow represents oxygen, cyan represents iron and the zone in between is tin segregated to the grain boundary and delineating the failure path. The role of grain boundary microchemistry in stress corrosion failure is described on page 663.

This issue also contains a review supplement devoted to textbooks published during the past eighteen months.

Contents continued overleaf

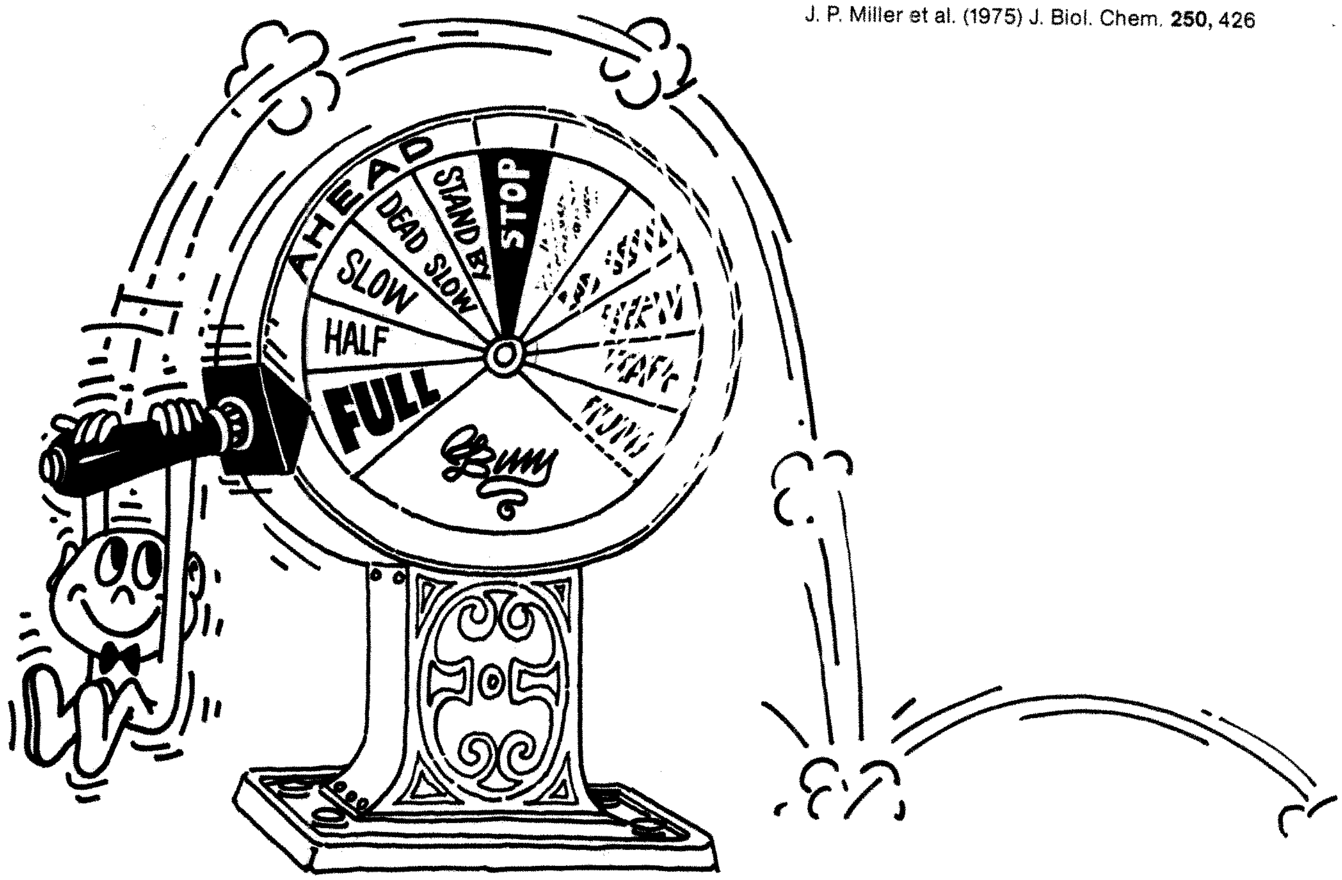
Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/08000C-03$01.00

Chlorophenylthio-cAMP

the potent activator of the cAMP-dependent protein kinase

J. P. Miller et al. (1975) J. Biol. Chem. **250**, 426



Boehringer Mannheim

the leading manufacturer of cyclophosphates and your supplier of:

adenosine-3′:5′-monophosphoric acid
8-bromoadenosine-3′:5′-monophosphoric acid
8-bromoguanosine-3′:5′-monophosphate
8-(4-Chlorophenylthio) adenosine-3′:5′-monophosphate
6-chloro-purine riboside-3′:5′-monophosphate
cytidine-3′:5′-monophosphate
2′-deoxy-adenosine-3′:5′-monophosphate
2′-deoxy-thymidine-3′:5′-monophosphate
N^6-2′-0-dibutyryl-adenosine-3′:5′-monophosphate
N^4-2′-0-dibutyryl-cytidine-3′:5′-monophosphate
N^2-2′-0-dibutyryl-guanosine-3′:5′-monophosphate
guanosine-3′:5′-monophosphate
guanosine-3′:5′-monophosphoric acid
inosine-3′:5′-monophosphoric acid
N^6-monobutyryl-adenosine-3′:5′-monophosphate
2′-0-succinyl-adenosine-3′:5′-monophosphoric acid
2′-0-succinyl-guanosine-3′:5′-monophosphoric acid
uridine-3′:5′-monophosphate

For details please contact your local representative of

mannheim boehringer

Boehringer Mannheim GmbH
Biochemica
P.O. Box 310120
D-6800 Mannheim
W. Germany

Circle No.36 on Reader Enquiry Card.

R.C. Thunell	**Cenozoic palaeotemperature changes and planktonic foraminiferal speciation**	670
W.E. Winner, C.L. Smith, G.W. Koch, H.A. Mooney, J.D. Bewley and H.R. Krouse	**Rates of emission of H_2S from plants and patterns of stable sulphur isotope fractionation**	672
P. Yodzis	**The stability of real ecosystems**	674
V.C. Tartter and K.C. Knowlton	**Perception of sign language from an array of 27 moving spots**	676
R.M. Gorczynski and E.J. Steele	**Simultaneous yet independent inheritance of somatically acquired tolerance to two distinct H-2 antigenic haplotype determinants in mice**	678
R.L. Coffman and I.L. Weissman	**B220: a B cell-specific member of the T200 glycoprotein family**	681
G. Dahl, R. Azarnia and R. Werner	**Induction of cell–cell channel formation by mRNA**	683
J.M. Dubois and M.F. Schneider	**Block of Na current and intramembrane charge movement in myelinated nerve fibres poisoned with a vegetable toxin**	685
T.K. Gartner, J.M. Gerrard, J.G. White and D.C. Williams	**Fibrinogen is the receptor for the endogenous lectin of human platelets**	688
D.G. Stephenson, I.R. Wendt and Q.G. Forrest	**Non-uniform ion distributions and electrical potentials in sarcoplasmic regions of skeletal muscle fibres**	690
A. Mangel, M. Fahim and C. van Breemen	**Rhythmic contractile activity of the *in vivo* rabbit aorta**	692
F. Jähnig	**No need for a new membrane model**	694
R.P. Silver, C.W. Finn, W.F. Vann, W. Aaronson, R. Schneerson, P.J. Kretschmer and C.F. Garon	**Molecular cloning of the K1 capsular polysaccharide genes of *E. coli***	696

MATTERS ARISING

G.R. Hervey and G. Tobin	**Brown adipose tissue and diet-induced thermogenesis**	699
N.J. Rothwell and M.J. Stock	**Reply to Hervey and Tobin**	699

TEXTBOOK SUPPLEMENT

	Self-consciousness and the book for students	701
D.J. Cove	**Mendel meets DNA: the teaching of genetics — genetics textbooks**	702
J.R.S. Fincham	**Genetics of Microbes (by Brian W. Bainbridge)**	704
B. Burnet	**Behavioral Genetics: A Primer (by R. Plomin *et al.*)**	704
D.H. Watson	**Introduction to Modern Virology, 2nd Edn (by S.B. Primrose and N.J. Dimmock)**	704
R. Whittenbury	**Microbiology, 3rd Edn (by G.A. Wistreich and M.D. Lechtman)**	705
Peter E. Long	**Introduction to Modern Mycology (by J.W. Deacon). Introduction to Fungi, 2nd Edn (by John Webster)**	705
J.G. Vaughan	**Plant Taxonomy and Biosystematics (by Clive A. Stace)**	706
James F. Sutcliffe	**Transport in Plants (by U. Lüttge and N. Higinbotham)**	706
Peter D. Moore	**Vegetation of the Earth, 2nd Edn (by Heinrich Walter). Plant Geography, 2nd Edn (by Martin C. Kellman)**	707
L.J. Audus	**Plant Hormones and Plant Development (by W.P. Jacobs) Biochemistry and Physiology of Plant Hormones (by T.C. Moore)**	707
Paul A. Colinvaux	**Biogeography: An Ecological and Evolutionary Approach, 3rd Edn (by C.B. Cox and P.D. Moore). Biogeography (by E.C. Pielou)**	708
I.J. Linn	**Understanding ecosystems — ecology textbooks**	708
C.M. Yonge	**Fundamentals of Aquatic Ecosystems (edited by R.S.K. Barnes and K.H. Mann). The Ecology of Streams and Rivers (by Colin R. Townsend)**	709
Aubrey Manning	**Introduction to Ethology (by K. Immelmann)**	710
W.C. McGrew	**Social Behaviour in Primates (by Neil Chalmers)**	710
Trevor W. Robbins	**More neuroscience — neuropsychology textbooks**	710
Stuart Sutherland	**More sight than sound — textbooks on perception**	711
Angus Gellatly	**Cognitive Psychology (by R.L. Solso). Cognitive Psychology and Its Implications (by J.R. Anderson)**	712
Stewart Evans	**Introduction to Insect Behavior (by M.D. Atkins)**	712
R.W. Thorp	**Entomology (by Cedric Gillott)**	713
Barry Cox	**Vertebrate Life (by W.N. McFarland *et al.*)**	713
John C.W. Cope	**Elements of Palaeontology (by Claude Babin)**	714
Alan Lord	**Microfossils (by M.D. Brasier)**	714
E.C. Cox and R.M. May	**Biological Science, 3rd Edn (by W.T. Keeton)**	714
P.T. Saunders	**Maths applied — mathematics and statistics for life scientists**	715
Joseph E. Coleman	**Biophysical Chemistry (by C.R. Cantor and P.R. Schimmel)**	716

ENVIRONMENTAL SCIENCE & TECHNOLOGY

ES&T

Environmental Science & Technology—The premier research journal in the environmental field

Environmental science continues to be one of the fastest growing fields. And ES&T has grown right along with it!

ES&T continues to give you the practical, hard facts you need on this science...covering research, techniques, feasibility, products and services.

Essential reading for environmental scientists both in the business and academic world...ES&T has increased its emphasis on peer-reviewed research dealing with water, air, and waste chemistry in addition to adding critical discussions of important environmental science issues—all relevant to understanding the management of our natural environment. Also included is analysis of governmental regulations, current environmental research lab activities, and much more!

Circle No. 19 on Reader Enquiry Card.

Use the coupon to order your subscription today

Environmental Science & Technology, American Chemical Society
1155 Sixteenth Street, N.W., Washington, D.C. 20036 1981

Yes, I want to keep up-to-date with ENVIRONMENTAL SCIENCE & TECHNOLOGY. Please enter my subscription as follows:

ACS Members*	U.S.	Foreign
1-Year	☐ $18.00	☐ $25.00
3-Years	☐ $43.00	☐ $64.00
Nonmembers Personal		
1-Year	☐ $21.00	☐ $28.00
3-Years	☐ $50.00	☐ $71.00
Nonmembers Institutional		
1-Year	☐ $72.00	☐ $79.00

Please specify: ☐ Hard copy, or ☐ Microfiche ☐ Payment enclosed (Payable to American Chemical Society) ☐ Bill me ☐ Bill company

Name________________

Title________________

Organization________________

Address ☐ Business ☐ Home________________

City________________

State________________ Zip________

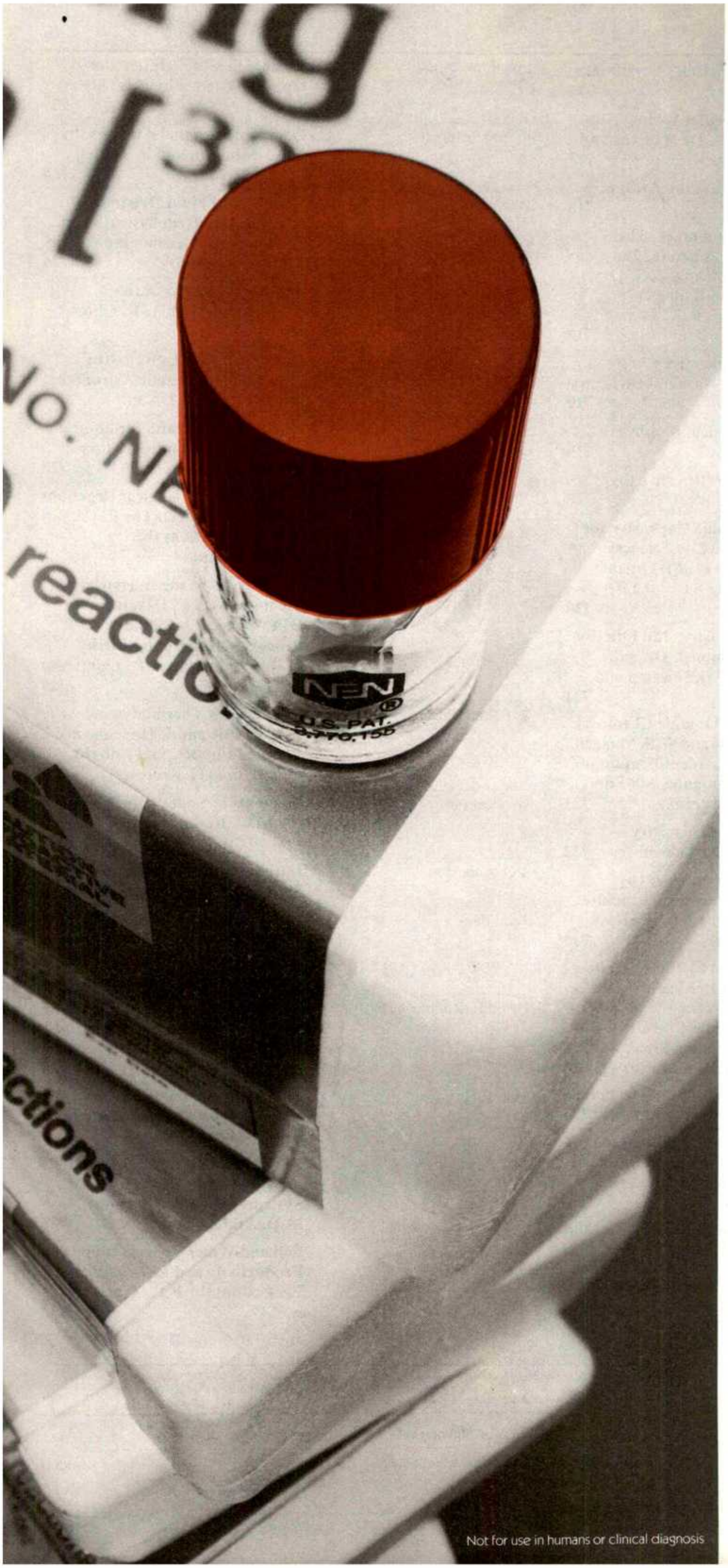
Allow 60 days for your first copy to be put in the mail. *ACS Member rates are for personal use only.

CALL TOLL FREE: 800-424-6747 1108H

David Saggerson	**Biochemical quartet — textbooks on biochemistry**	**716**
John Edwards	**Cell and Molecular Biology, 7th Edn (by E.D.P. De Robertis and E.M.F. De Robertis)**	**717**
K. Murray	**DNA Replication (by Arthur Kornberg)**	**718**
Dennis Chapman	**Biological Membranes: Their Structure and Function, 2nd Edn (by R. Harrison and G.G. Lunt). Introduction to Biological Membranes (by M.K. Jain and R.C. Wagner)**	**718**
Arden P. Zipp	**Introductory chemistry in the market place — general chemistry for beginners**	**719**
R.G. Gosling	**Topics in Classical Biophysics (by H.J. Metcalf)**	**719**
J.W.T. Dickerson	**Nutrition: in Perspective (by Patricia A. Kreutler)**	**720**
Peter Garratt	**Essential Organic Chemistry for Students of the Life Sciences (by A.P. Ryles *et al.*). Organic Chemistry: The Basis of Life (by B. Miller)**	**720**
John Mann	**Organic Chemistry, 2nd Edn (by T.W.G. Solomons). Organic Chemistry (by D.S. Kemp and F. Vellaccio)**	**721**
Paul Schatz	**Experimental Organic Chemistry (by M.P. Doyle and W.S. Mungall). The Systematic Identification of Organic Compounds, 6th Edn (by R.L. Shriner *et al.*)**	**721**
Charles A McAuliffe	**Joys of general chemistry — more introductory chemistry**	**722**
A.J. Stone	**Orbitals and Symmetry (by D.S. Urch). Chemical Structure and Bonding (by R.L. DeKock and H.B. Gray)**	**723**
Roy M. Harrison	**Environmental Chemistry: The Earth–Air–Water Factory (by R.W. Raiswell *et al.*). Experiments in Environmental Chemistry: A Laboratory Manual (by P.D. Vowles and D.W. Connell)**	**723**
John Barkham	**Environmental what? — textbooks on environmental management**	**723**
C. Kidson	**Process in Geomorphology (edited by C. Embleton and J. Thornes). Geomorphological Processes (by E. Derbyshire *et al.*)**	**724**
Bruce Sellwood	**Origin of Sedimentary Rocks, 2nd Edn (by H. Blatt *et al.*)**	**725**
D. Flinn	**Principles of Mineral Behaviour (by A. Putnis and J.D.C. McConnell)**	**725**
Robert N. Colwell	**Remote Sensing: Optics and Optical Systems (by P.N. Slater)**	**725**
Freeman Gilbert	**Geophysical Signal Analysis (by E.A. Robinson and S. Treitel)**	**726**
D. Anderson	**Geophysical Fluid Dynamics (by J. Pedlosky). Geophysical Fluid Dynamics for Oceanographers (by J.J. von Schwind)**	**726**
John Green	**Atmospheric Physics (by J.V. Iribarne and H.-R. Cho)**	**727**
Stephen P. Maran	**Discovering the Universe (by C.E. Long). Dynamic Astronomy, 3rd Edn (by R.T. Dixon)**	**727**
R.B. Jones	**Quantum mechanics for librarians — textbooks on introductory quantum physics**	**728**
James Lowe	**Introduction to Nuclear Reactions (by G.R. Satchler). The Physics of Nuclear Reactions (by W.M. Gibson)**	**728**
R.J. Tayler	**The Physics of the Interstellar Medium (by J.E. Dyson and D.A. Williams). Radiative Processes in Astrophysics (by G.B. Rybicki and A.P. Lightman)**	**729**
P.T. Landsberg	**Principles of Thermodynamics (by J.A. Beattie and I. Oppenheim). Thermal Physics, 2nd Edn (by C. Kittel and H. Kroemer)**	**729**
W. Cochran	**Elements of Solid State Physics (by M.N. Rudden and J. Wilson)**	**730**
J.C. Anderson	**Introduction to Applied Solid State Physics (by R. Dalven)**	**730**
John Pain	**Electricity and Magnetism, 3rd Edn (by W.J. Duffin). Fundamentals of the Theory of Electricity (by I.E. Tamm)**	**730**
David L. Rimmer	**Soil Physics (by T.J. Marshall and J.W. Holmes). Applied Soil Physics (by R.J. Hanks and G.L. Ashcroft)**	**731**
Milo I. Harpstead	**Soils: Their Formation, Classification and Distribution (by E.A. FitzPatrick)**	**731**
J.C. Bowman	**Agricultural Ecology: An Analysis of World Food Production Systems (by G.W. Cox and M.D. Atkins)**	**732**
Ronald F. Hensler	**Soil and Water Conservation (For Productivity and Environmental Protection) (by F.R. Troeh *et al.*)**	**732**

MISCELLANY

Erratum	**698**
Books received	**733**
New on the market	**xxx**
Classified advertising	**xxxi**



Not for use in humans or clinical diagnosis

Another first for DNA-RNA studies

With the introduction of our new 3' End Labeling System, we now offer six systems, plus special accommodation packaging, to make your DNA-RNA studies faster, surer, and more reproducible.

Only from NEN

nature

Volume 289 No.5800 26 February 1981

Darwin's death in South Kensington 735

New Pentagon rules on overseas students: Secrecy threat to sensitive research plans 736

British gene company opens books 736

Celltech deal with MRC 737

Polish universities: Sit-in complete 737

US science budget: Reagan's way 737

Threat to solar mission 738

UK nuclear energy: CEGB sheepish 738

Authority more critical 739

UK science research: On the move 739

Community research: Project sharing 739

US Cabinet appointments: Conservatives in 740

UK medical research: Money to spend 740

High-energy physics: China backs off 740

Soviet space research: Signs of strain 741

Science and government: Lords look now 741

Questions to answer 741

Correspondence: Genes and race (J. Maynard Smith); Museum debate (S. J. Gould); Last word? (L. B. Halstead); Badger controversy (J. R. Flowerdew) 742, 831

NEWS AND VIEWS

Models of speciation — the evidence from Drosophila: J.S. Jones assesses natural selection and 'genetic revolution' as explanations for the origin of species 743

Last supper of two Indian Triassic reptiles? Barry Cox describes a new genus of reptile 745

Is the Universe expanding? David A. Hanes on the implications of recent statistical studies 745

Parasitology in Venezula: Neil Brown and Win Gutteridge describe current research 746

Weighing neutrinos: a small wait for a small weight? Frank E. Close describes a new approach 747

Woodchuck, squirrel and duck hepatitis viruses: Arie J. Zuckerman compares hepatitis viruses 748

Hominoid habitats of the Miocene: Peter Andrews on how our ancestors lived 749

Variability of the Earth's rotation: Frank D. Stacey on a recent meeting 750

REVIEW ARTICLE

C. Yanofsky **Attenuation in the control of expression of bacterial operons** 751

ARTICLES

R.W. Porcas, R.S. Booth, I.W.A. Browne, D. Walsh and P.N. Wilkinson **VLBI structures of the images of the double QSO 0957+561** 758

J.V. Smith **Halogen and phosphorus storage in the Earth** 762

R.M. Anderson, H.C. Jackson, R.M. May and A.M. Smith **Population dynamics of fox rabies in Europe** 765

T. Blundell, P. Lindley, L. Miller, D. Moss, C. Slingsby, I. Tickle, B. Turnell and G. Wistow **The molecular structure and stability of the eye lens: X-ray analysis of γ-crystallin II** 771

LETTERS

D. Toussaint and F. Wilczek **Constraints on heavy neutrinos** 777

U.R. Rao, T.K. Alex, V.S. Iyengar, K. Kasturirangan, T.M.K. Marar, R.S. Mathur and D.P. Sharma **IR observations of the solar corona — a ring around the Sun?** 779

J.F. McCann, M.S. Kazacos and D. Haneman **Temperature dependence for the power outputs of *n*-CdSe liquid junction cells** 780

R. von Ballmoos and W.M. Meier **Zoned aluminium distribution in synthetic zeolite ZSM-5** 782

S. Blake **Volcanism and the dynamics of open magma chambers** 783

M. Wyss and M. Baer **Seismic quiescence in the Western Hellenic Arc may foreshadow large earthquakes** 785

R. Batiza and C.G. Chase **Miocene spreading centre south of Isla Guadalupe** 787

Contents continued overleaf

Cover: An electron micrograph of a gelatin polymer network. Comparison with network structures in gels of rod-like polypeptides is described on page 813.

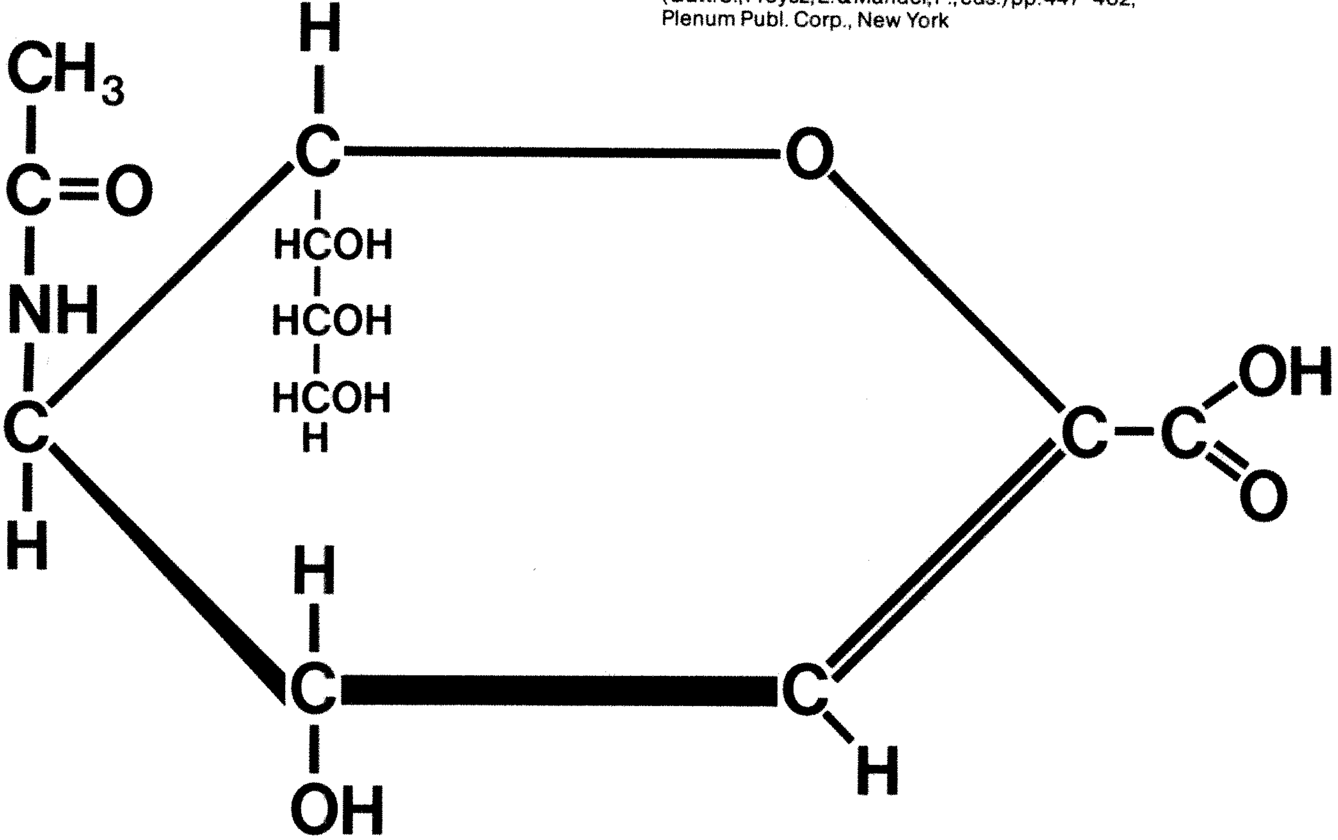
M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

2,3-Dehydro-2-deoxy-N-acetylneuraminic acid

the competitive inhibitor of neuraminidases

Veh, R. W. & Schauer, R. (1978) in Enzymes of Lipid Metabolism (Gatt. S., Freysz, L. & Mandel, P., eds.) pp. 447–462, Plenum Publ. Corp., New York



is now available with

BOEHRINGER MANNHEIM

your supplier of:

N-acetylneuraminic acid	Cat. No. 101 923
N-acetylneuraminosyl-D-lactose	Cat. No. 161 721
N-acylneuraminic acid glycoproteid	Cat. No. 102 032
2,3-dehydro-2-deoxy-N-acetylneuraminic acid	Cat. No. 528 544
2-(3'-methoxyphenyl)-N-acetyl-D-neuraminic acid	Cat. No. 174 696
neuraminic acid-ß-methyl glycoside	Cat. No. 107 603
neuraminidase from Arthrobacter ureafaciens	Cat. No. 269 611
neuraminidase from Clostridium perfringens	Cat. No. 107 590

For details please contact your local representative or

mannheim boehringer

Boehringer Mannheim GmbH
Biochemica
P.O. Box 310120
D-6800 Mannheim
W. Germany

Circle No.27 on Reader Enquiry Card.

LETTERS

M.E. Cox and G.M. McMurtry **Vertical distribution of mercury in sediments from the East Pacific Rise** 789

Z. Zhongying **Precambrian microfossils from the Sinian of South China** 792

J.H. Lawton and M.P. Hassell **Asymmetrical competition in insects** 793

A.S.C. Medcalf and P.D. Lawley **Time course of O^6-methylguanine removal from DNA of *N*-methyl-*N*-nitrosourea-treated human fibroblasts** 796

C.W. Stackpole **Distinct lung-colonizing and lung-metastasizing cell populations in B16 mouse melanoma** 798

E. Mitrani and H. Eyal-Giladi **Hypoblastic cells can form a disk inducing an embryonic axis in chick epiblast** 800

L.-M. Ching and R.G. Miller **Generation of cytotoxic T-lymphocyte precursor cells in T-cell colonies grown *in vitro*** 802

S.C. Riley, E.J. Brock and W.M. Kuehl **Induction of light chain expression in a pre-B cell line by fusion to myeloma cells** 804

D. Holowka, C. Gitler, T. Bercovici and H. Metzger **Reaction of 5-iodonaphthyl-1-nitrene with the IgE receptor on normal and tumour mast cells** 806

C.J. Kenyon and G.C. Walker **Expression of the *E. coli uvrA* gene is inducible** 808

R. McKay and D. DiMaio **Binding of an SV40 T antigen-related protein to the DNA of SV40 regulatory mutants** 810

K. Tohyama and W.G. Miller **Network structure in gels of rod-like polypeptides** 813

D.W. Banner, C. Nave and D.A. Marvin **Structure of the protein and DNA in fd filamentous bacterial virus** 814

M.A. Viswamitra, O. Kennard, W.B.T. Cruse, E. Egert, G.M. Sheldrick, P.G. Jones, M.J. Waring, L.P.G. Wakelin and R.K. Olsen **Structure of TANDEM and its implication for bifunctional intercalation into DNA** 817

J.S. Pober and J.L. Strominger **Transglutaminase modifies the carboxy-terminal region of HLA-A and -B antigens** 819

H. Meier, L. Buchs A.J. Buchala and T. Homewood **(1→3)-β-D-Glucan (callose) is a probable intermediate in biosynthesis of cellulose of cotton fibres** 821

MATTERS ARISING

M.H. Wolpoff; **Allez Neanderthal** 823

C.B. Stringer, R.G. Kruszynski and R.M. Jacobi 823

A. ApSimon **Reply to Wolpoff** 823
Reply to Stringer, Kruszynski and Jacobi 824

R.D. Montgomerie **Why do jackals help their parents?** 824

P.D. Moehlman **Reply to Montgomerie** 825

M.A. Georgeson and M.G. Harris **Size constancy does not fail below half a degree** 826

BOOK REVIEWS

Eric Ashby **Societal Risk Assessment: How Safe is Safe Enough? (edited by R.C. Schwing and W.A. Albers Jr)** 827

R.K.S. Wood **Fungus Diseases of Tropical Crops (by Paul Holliday)** 827

Reginald Newell **Climate and Weather in the Tropics (by H. Riehl)** 828

J.P. Goddard **Transfer RNA: Biological Aspects (edited by D. Söll, J.N. Abelson and P.R. Schimmel)** 828

Clive Ellory **Epithelial Transport in the Lower Vertebrates (edited by B. Lahlou)** 829

Robert Hudson **A Complete Checklist of the Birds of the World (by Richard Howard and Alick Moore)** 829

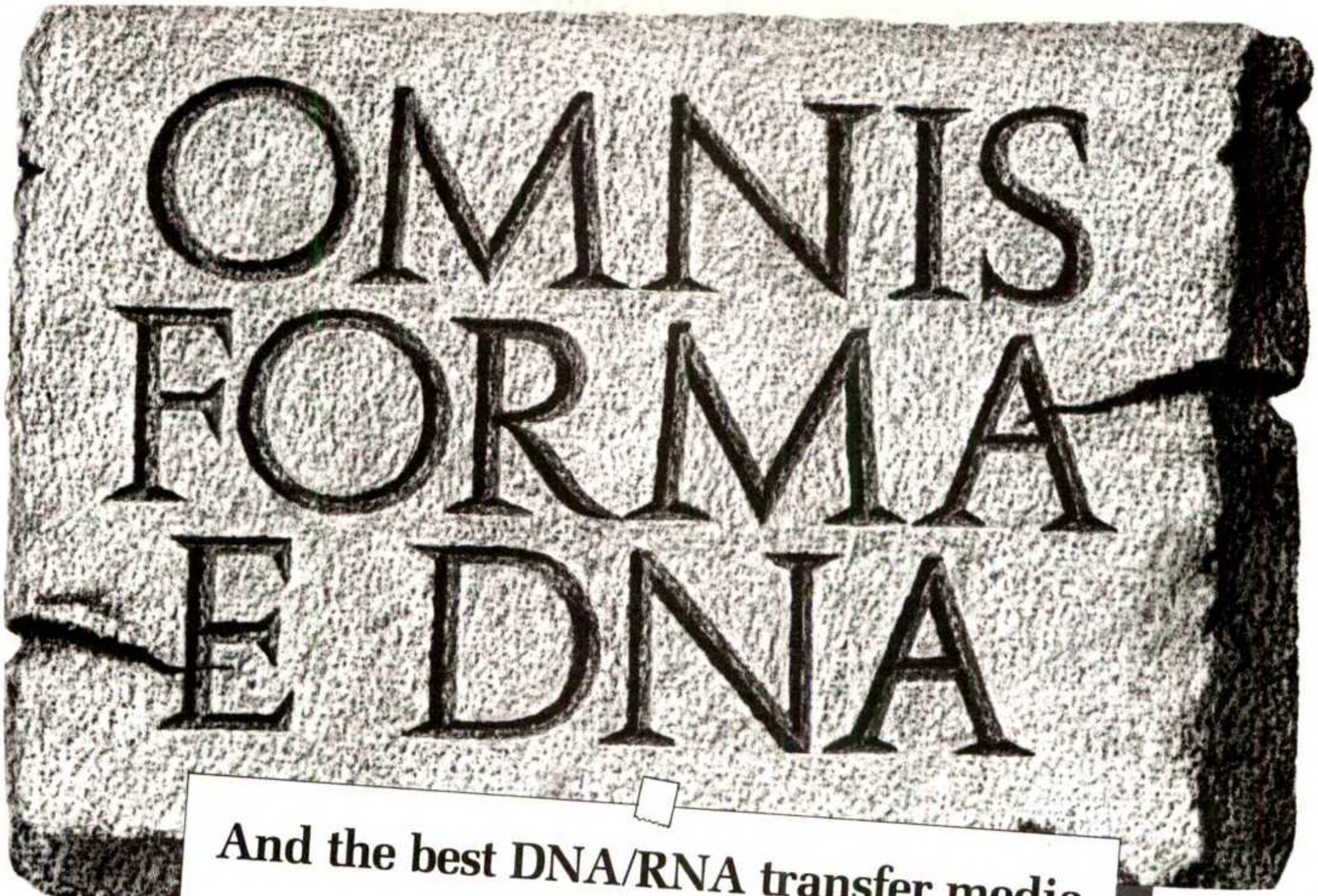
S.D. Smith **Physics of Semiconductor Laser Devices (by G.H.B. Thompson)** 830

P.C. Crofts **Infrared Characteristic Group Frequencies (G. Socrates)** 830

MISCELLANY

100 years ago 744
Books received 832
Classified advertising xxvii

PRODUCT REVIEW This week's product review (page xxi) covers microbiological products



And the best DNA/RNA transfer media come from S&S.

S&S transfer media are established as the products of choice for "Southerns," "Northerns," "Westerns," "shotgunning," "plaque lifts" and "blitz-blotting"—because they do the job.

Pure nitrocellulose—the proven standard

Because S&S nitrocellulose is *pure* nitrocellulose, it effects a high degree of DNA transfer from gels, bacterial and plaque lawns. Maximum levels of hybridization can be achieved, resulting in more intense bands or spots. S&S nitrocellulose is available in whatever form you need: 12-0.025 μm pore size, gridded, in circles, squares, strips, rolls or full sheets.

Transa-Bind™—universally accepted medium for "Northern" transfers

Schleicher & Schuell Transa-Bind medium has found wide use in DNA, RNA and protein transfers because the fragments are covalently bound to the medium. Available in NBM (nitro) form and ABM (amino) form to suit your needs. Packaged in nitrogen-filled opaque bags to extend shelf life.

New S&S products in development for genetic researchers

S&S enjoys an earned reputation for working with researchers in developing new products to meet their specific needs. Be on the alert for new S&S transfer media soon to be introduced. Contact S&S for any information or technical help you might need.

See us at Booths #428-429 at ASM

Schleicher & Schuell

Schleicher & Schuell, Inc.
Keene, New Hampshire 03431
Schleicher & Schüll GmbH, D-3354.
Dassel, West Germany
Schleicher & Schüll AG, CH-8714
Feldbach ZH, Switzerland

Circle No. 17 on Reader Enquiry Card.

Latin, "Omnis forma e DNA," — "All forms (of life) from DNA."

nature

1/8 January 1981

Signs of grace (and opposite) for 1981

All new years are important. This, 1981, will be crucially so. By next December, it should be clear how several influential governments are tackling the problems which now confront them. For the time being, the most absorbing question is whether Mr Ronald Reagan will live up to his election promises — and threats. Before the year is up, however, there should be a new government in the Soviet Union, Even when the fear that change means danger is as deeply ingrained as in Moscow, it is unthinkable that Mr Brezhnev will be kept on as president much longer. In Western Europe, the chances are that the same old governments will have to battle with the same old problems, but with a sense of mounting despair. If the problems of economic stagnation coupled with inflation appear as obdurate by the end of the year as they seem now, people will begin to behave as if these problems have come to stay. So what are the signs to look for in the months ahead?

Economic problems matter because they determine the resources available for existing, cherished, institutions and for innovation. The industrialized democracies now all suffer to some degree from the disease pioneered first by Italy and then by Britain — rapid inflation born of democratic voters' convictions that living standards must inexorably rise, even when productivity is falling or when resources are transferred in substantial amounts to OPEC oil producers. The consequences are all too clear — industrial investment and research and development are starved of resources, productivity falls still further, the pool of wealth for sharing out diminishes and the dole queues lengthen. One of the first tests of Mr Reagan's temperament in office rather than on the hustings will be the fine print that will no doubt accompany his promised tax-cut. The promise is too fresh in people's minds to be forgotten, but a tax-cut unrequited by reduced federal expenditure would be inflationary. So will the offsetting cuts be intelligent (housing subsidies, pointless environmental protection, excessive solar energy research), stupid (education, research and development) or — horror — non-existent? Other governments in industrialized countries are likely to take their cue from Washington. If Mr Reagan ducks the economic issues, many other governments will know that there is no point in their struggling separately against inflation.

The Soviet economic problem, outwardly different, is as worrying — and not exclusively for Soviet citizens. In the whole Soviet bloc, there is less wealth to share than the managers of the economy would like. Inflation is, by definition, not a problem. Whatever people's hopes, their shares of what is to be shared are determined by the system of allocation. One consequence is the grinding deprivation of much of the Soviet system, partly the cause of troubles such as those with which Poland has recently been beset. The question which the anonymous electors of Mr Brezhnev's successor will have to face is whether the new Soviet government should risk some relaxation of the domestic system for the benefits that would ultimately accrue in improved economic performance. Probably they will shrink from taking such a step on the grounds that the risks of relaxation would be immediate and the benefits delayed. That will be bad news both for citizens of the Soviet bloc and for the rest of us. The needless economic privations of Eastern Europe and beyond are not simply cruel to those who bear the brunt of them but part of the cause of continuing trouble between East and West.

The irony, both in the West and the East, is that the ingredients of continuing technical change and economic improvement are clearly apparent. The cost of petroleum has become a burden, but nuclear power stations are partially a substitute. The new computer technologies, wildly oversimplified though they may have been, are gathering dust for lack of investment capital and sufficiently skilled manpower. Communications technology, physical and electronic, is held back by lack of capital but also by regulatory bumbledom. Agriculture, surprisingly the great success of the industrialized West in recent decades, languishes elsewhere. The outstanding challenge for most governments in 1981 will be to make a fresh start on using these ingredients of prosperity to enable their people to feed, clothe and house themselves decently, and to have something left over to do what they like with. For many Western governments, but especially for the American, this will require something of a battle with the powerful lobbies which have grown up in the past decade and which appear, after all, to think that there is such a thing as a free lunch — the benefits of technology but no generating plants within sight or the benefits of new drugs but no experiments with animals please. Mr Reagan's reputation suggests that he might throw himself cheerfully into such a fight, but only time will tell.

Two other issues will be bellwethers for the coming year, with policies on defence and arms control most vividly in people's minds. Mr Reagan seems to be committed to winning the arms race, whatever that may mean. Nobody, not even the Soviet Union, will be surprised if the Pentagon's budget now grows again. People in Europe will bite their nails. The danger is that one side or the other will be tempted to check whether it is really winning by pressing all the buttons. The chances are, however, that the arms race of the 1980s will not be very different from that of the 1950s — the modern equivalent of sabre rattling, a diversion of resources from more constructive purposes but adventitiously a source of further technological innovation. What matters much more, especially where Mr Reagan is concerned, is whether he will have the wit and the courage (for a Republican) to seize the opportunity in arms control with which he has been presented. The new administration is committed to reopening Salt II, but Mr Reagan is not the out-and-out abrogator who began the election campaign. As Inauguration Day approaches, he will be made increasingly aware of the advantages, philosophical and political, from balancing his part in the arms race with an open-minded view of what package might replace Salt II. The Comprehensive Test-Ban Treaty, nearly ready for signature though it is, is probably too much of a pill for the Republican Senate and General Alexander Haig to swallow. There is more room for manoeuvre on what are called theatre nuclear weapons — the missiles that must worry the West Germans but also, no doubt, Soviet military planners as well.

A less obvious sign of the way the wind is blowing in 1981 will be the way in which governments, but especially those of industrialized countries, shape in their dealings with the developing world. Mr Reagan's reputation is not encouraging but, to be fair, his position on the recommendations of the Brandt Report (as on many other issues) is not all that explicit. Again, the new president has the appearance of a conservative who may be pushed in liberal directions by events or even arguments. The obvious difficulty is that the months ahead are likely to see still more occasions on which developing countries act uppishly towards the United States (provoking pressure to put them in their place), but that there is no powerful political constituency for the decently self-interested policy of adequate assistance, technical and financial. In the long run, however, there will be no point in beating inflation or surviving the arms race ahead if the world that is left remains as dangerously divided into rich and poor.

0028-0836/81/010001-01$01.00

Beating a retreat from Rothschild

British governments have long experience of saying one thing and doing another, and the skills thus acquired are now being applied to the management of civil science. Less than ten years have passed since a previous Conservative government (Mr Edward Heath's) embraced Lord Rothschild's then radical notion that government departments seeking benefits from applied research should pay for it out of their own budgets. For most of the 1970s, the British government was devoted to this principle, with the result that government departments (called "customers") and research councils (called "contractors") have been locked in almost perpetual negotiation about the terms on which specific research projects should be carried out. When the Rothschild era began, the research councils affected were deeply offended, and for a time it seemed as if the then president of the Royal Society, Sir Alan Hodgkin, would join them on the barricades. In the event, it seems that the anti-Rothschild forces need not have bothered. They could have counted on the unwillingness of government departments to change their ways.

A few weeks ago, it became clear that the Medical Research Council had won its decade-long battle with the Department of Health, and had won back the share of its budget filched by the department with effect from 1973 (see *Nature* 23 October). Now, the Agricultural Research Council's relationship with its principal "customer", the Ministry of Agriculture, Fisheries and Food, is being changed but in a different direction — the ministry seems likely to have a more direct say in what the council does (see *Nature* 11 December).

This turn of events has an obscure origin — a report of the Public Accounts Committee of the House of Commons published in July 1979 dealing with the cost of storing skimmed milk and potatoes as well as with the affairs of the two research councils. The proceedings, nevertheless, have an arcane interest. Dr James Gowans, the secretary of the Medical Research Council, gave the committee a winsome account of the importance of good luck in the conduct of medical research. The steps in medical research in the past twenty years that "have made the greatest impact. . . have all been made by accident", and as a consequence there is nothing that a research council can do but back "good ideas and good people". Customer-directed research is, in other words, impossible.

The committee, known to the popular newspapers in Britain as the "watchdog" over public expenditure, softened to putty, swallowed the line taken by Dr Gowans and the Department of Health that it was "impracticable" for the ministry to control the research it had commissioned, and recommended that for medical research the government should abandon the Rothschild principle. Both parties have lost no time, and the arrangement has indeed been unscrambled.

With agriculture, similar arguments appear to have led the Public Accounts Committee to the opposite conclusion. The committee seems to have been bamboozled into believing that the Agricultural Research Council supported research intended to yield practical benefits (plant-breeding, for example) with funds left untouched by the Rothschild transfer to the ministry. The committee, evidently ignorant of the council's terms of reference, kept asking why. And why should there have been so little change in the pattern of the council's work since 1972? Surely the only explanation is that the ministry's influence was too remote, and the only possible remedy that the ministry should have a more direct say in what the council does? The ministry has leapt at the chance. It may not be long before the Agricultural Research Council is indistinguishable from a government department.

This is an odd way in which to tackle an issue of principle. The Public Accounts Committee of the House of Commons, for all its virtues, is neither experienced nor competent in fields like this. Its cursory inquiry into the workings of the Rothschild principle was hardly searching enough to give the two ministries the licence that they have so eagerly assumed. The conclusion may be correct that the Rothschild arrangement has not functioned well in medical and agricultural research, but the explanations given are spurious. Dr Gowans's doctrine of investment in "good ideas and good people" is a necessary but not a sufficient condition for success — and applies in other fields than medical research. And if neither the medical nor the agricultural research programmes have been changed in pattern since 1972, the explanation may simply be the indolence, incompetence or neglect of the ministries concerned. The unhappy tale of how the Department of Health set out to act as a customer by setting up a network of committees whose meetings were as often cancelled as held has been vividly documented in Professor Maurice Kogan's study (with Nancy Korman and Mary Henkel) *Government's commissioning of research* (Department of Government, Brunel University). The wonder is not that the system failed to work, but that its designers thought it might.

Before the rot goes too far, the British government must decide what it is trying to accomplish. Lackadaisical support for supposedly autonomous research councils is demoralizing, and probably wasteful of public money as well. The plain truth is that three of the research councils (agriculture, medicine and natural environment) are geared to basic research relevant to the practice of technology (*pace* physicians). Their autonomy is an historical accident, going back to the First World War and Lord Haldane's doctrine that government departments should not be allowed to control basic research because they would not understand it. Government departments may not be much better now, but times have changed. The Advisory Board for the Research Councils should be the forum in which the research councils make their priorities explicit. In practice, the board seems to be yet another device for making sure that no research council is disappointed in its share of the collective budget. How could it be otherwise, when the chief executives of the councils dominate the proceedings?

The truth is that the five research councils are different animals. The Science Research Council is an essential adjunct of higher education. The Social Science Research Council was set up both to encourage academic social science and to apply the social sciences to the solution of practical problems; it has been only a modest success, but these are early days. All five research councils, together with the Royal Society and the scientific museums, share in the annual cake-cutting ceremonies of the advisory board. (There is even talk that the British Academy, responsible for research in the humanities, should be included.) The result is that the question of how much of the civil research budget should be spent in or on universities, and how much on basic research relevant to technology, is never openly considered.

So what should the British government do? The simple response is to say that there should be yet another inquiry into the organization of civil science. The interval since Rothschild has been longer than any other period since the Second World War when some mammoth reorganization has not taken place. Such a decision, however, would be disastrous. The outcome would be another reorganization which, in the end, would be eroded by the unwillingness of those concerned to change their ways. The solution, which is being forced on the Thatcher government by its unthinking retreat from Rothschild, is a return to the arrangements of the early 1960s, when civil science was the responsibility of a part-time minister. The immediate benefit would be that there would be machinery in being for making policy as the need arose; as things are, nothing is dicided between agonized reappraisals. Such a development would have the extra benefit of separating the administration of civil science from that of education (for which Lord Todd was asking in his valedictory anniversary address to the Royal Society in November and will now no doubt be urging on the House of Lords Select Committee on Science and Technology). How much longer will it be before the government recognizes that its neglect of the past few months now predicates such a course?

0028-0836/81/010002-01$01.00

AAAS plans pressure on education

Dismay at proposed NSF budget cut

Toronto

The American Association for the Advancement of Science (AAAS) has decided to devote a substantial proportion of its efforts over the next decade to reversing what the US Department of Education has recently described as a growing tide of "scientific illiteracy".

Concerned about the growing gap between scientists and technologists on the one hand and the public and congressional decision-makers on the other, the association plans a series of programmes aimed, in particular, at raising the general level of scientific and mathematics education in the nation's primary and secondary schools.

The announcement was made during the AAAS's 147th annual meeting in Toronto this week. On Sunday, the association's board passed a resolution pledging the AAAS to work with its affiliated science and engineering societies to reverse what it called a "damaging decline" of science and engineering education in the United States.

Mr William Carey, the association's executive secretary, later said that the Carter Administration's apparent intent to cut back expenditure on science education by the National Science Foundation (NSF) was "a severe malfunctioning of the bureacracy".

This cut was occurring, he said, only two months after the NSF and the Education Department had delivered a report to the President which concluded that there had been a 15-year decline in the national commitment to excellence and international leadership in science, mathematics and engineering.

Mr Carey added that as a result of the Administration's decision to absorb cuts in the NSF budget by focusing in particular on its science education programmes, under the terms of the budget proposals which the President will present to Congress next week, science education would be reduced to only 7.5 per cent of NSF's efforts — the lowest figure for 30 years.

What concerns AAAS is not so much the quantity or quality of university science graduates — the department's report found little problem here, and suggested that where there were shortages or deficiences, these would probably be corrected by market forces — but evidence of declining educational standards in schools.

Dr Allan Bromley, professor of physics at Yale and next year's president of AAAS, said that it had been a "dismal mistake" to leave things in the hands of "professional educators". He cited, in particular, national surveys of educational achievement which demonstrated a general decline in scientific and mathematical abilities among high school students, as well as widespread evidence of the remedial action which universities and colleges were having to take with undergraduates enrolled in science classes.

"A whole generation of students is being short-changed" Dr Bromley said. "In physics, if you look at undergraduate curricula across the country, the first year is being increasingly devoted to teaching subjects that you would not have had to worry about 10 or 15 years ago."

Other AAAS officers were slightly less harsh in their criticisms. Current president Dr Frederick Mosteller of the Harvard School of Public Health pointed out that many subjects, such as calculus, were now widely taught in schools where they had been missing a generation ago. Retiring president Dr Kenneth Boulding spoke up in support of less conventional teaching methods, arguing that "education is becoming less important as a means of learning".

However, what Dr Bromley described as a "severe crisis" in science education has spurred the AAAS into convening a conference of the heads of its affiliated societies to look at the health and priority needs of science and engineering education in the United States in the 1980s — and to make educational excellence in these areas a major theme of next year's meeting.

The association also intends to use its new monthly magazine, now called *Science 81*, to produce teaching materials for use in secondary schools, and it announced that Dr F. James Rutherford, previously associate director of NSF responsible for science education, and at present assistant secretary of the Department of Education, will join AAAS as adviser on science education to the board of directors after leaving office on 20 January.

"The AAAS cannot solve the problems on its own, but if we don't get busy, things will get worse" said Dr Mosteller, adding that both federal agencies and private industry would probably be approached to support the association's proposed activities.

David Dickson

Slow progress on Greek reforms

As Greece enters the European Community (from 1 January) the universities are still in the throes of reorganization, with student unrest at some institutions. But the new Science Research and Technology Agency is in good shape, with two years of constructive work behind it, although the threat of its becoming ensnared in Greek bureacracy remains. The hope is that the influence of the community will strengthen the recent tendency to reform.

The government's first attempt at university reform (see *Nature* 2 February 1978) met with strong opposition and the enforcement of some of the provisions of the new Law 815 had to be postponed. A committee was set up to make new proposals and these have now been presented to the government. The success of the committee, representing the fourteen university-level institutions in Greece, owes much to the way in which its chairman, Professor F. Mitsis, enjoys the confidence of all sides.

Much deadwood was painlessly removed from the academic profession by the provisions in Law 815 for voluntary retirement — in one department of the University of Athens only three out of twelve are still in post. Now there are many openings to be filled using new selection procedures.

The new draft bill would abolish the unpopular system of "chairs", in which professors were surrounded by a mixed bag of assistants. Instead subject "sectors" would be introduced, each consisting of at least seven academics at three grades: lecturers, associate professors and professors. The short-term appointment of assistants, usually working for higher degrees, would also be allowed. Undergraduates are granted 20 per cent participation in all university bodies under the proposed law.

Examination regulations are being drafted separately, amid protests from student associations that as long as general conditions in the universities leave so much to be desired, students cannot be expected to attain high standards. At the Technological University at Athens there was a prolonged sit-in to protest against the system of semesters, considered unduly onerous by many students. And at the University of Athens chemistry students have been striking against the alleged harshness of a particular professor.

Greece's newest university, in Crete, remains in a critical state. Facilities are makeshift and most of the teaching depends on visiting lecturers from abroad. Permanent senior appointments are being blocked by the Athens-based administration committee, perhaps under the influence of the old establishment. Fotis Kafatos of Harvard is no longer a member.

For research, the outlook is more cheerful. More than two years ago the government invited Dr George Argyropoulos back from the United States to head the new Science Research and Technology Agency (see *Nature* 23 February 1978). The main achievement of the agency has been the introduction of

0028-0836/81/010003-02$01.00

project grants on the pattern of the British research councils, instead of state subsidies for the overall budgets of institutions.

Following wide consultations, the agency prepared a list of nine priority areas of research, approved by a ministerial committee in June 1979. By October 1980 436 grant applications had been received and some 70 approved projects were already under way. Applications are being dealt with speedily, the process taking an average of seven months including the time required for three referees (at least one from abroad) to report.

Other achievements include eighteen international agreements for joint projects. Solar energy plants are to be built with Germany and France, and a geothermal plant constructed on the island of Melos in cooperation with the European Economic Community.

There are some clouds on the horizon, however. The law setting up the Science Research and Technology Agency sought to exempt it from the notoriously bureacratic public accounts regulations, but the Ministry of Finance nevertheless managed to block payments of grants for "technical reasons". The Committee of Ministers which oversees the work of the agency has reaffirmed its confidence in it. One hopes that it will now be allowed to function unhindered. **E. M. Pantelouris**

UK space policy

Year of decision

Last year was a busy one for makers of British space policy. By November, a committee of the Central Policy Review Staff, the think tank, had submitted to the Cabinet its study of Britain's efforts in space applications; and an interdepartmental committee, under the Department of Industry, had been created to coordinate space policy more effectively and to discuss the think tank's deliberations. The Home Office was also busy preparing a report on direct broadcast television by satellite.

The sudden interest in space seems to have been stimulated by the fear that Britain might miss out on the profits that could be made from selling space technologies, especially telecommunications satellites. The think tank's report is to remain confidential for commercial reasons. The gist of the recommendations is that there is a demand for space applications satellites which British industry should be encouraged to meet.

One important question is whether Britain should try to build up its industry alone or whether it should continue, perhaps at an increased level, in the space applications programmes of the European Space Agency (ESA) which France and Germany, in particular, have used more effectively than Britain to boost their own industries. The most likely outcome is that Britain will continue at more or less the same level in ESA's telecommunications programmes, but that greater efforts will be made to transfer the results of ESA's research and development to industry.

The government is unlikely to rush to pour money into the industry, seeking rather to encourage private investment. A central issue in the telecommunications field will be the government's attitude to the monopolies held in television broadcasting by the broadcasting authorities and in satellite communications by the telecommunications division of the Post Office, British Telecom. A bill to dilute British Telecom's monopoly is now before Parliament, and could be used by the Department of Industry to liberalize access to satellite communications, leaving the way open for private operators, in particular, of small satellites for business communications. If greater incentives are given to satellite operators, the next question will be the ability of British Aerospace (which not everyone is convinced could withstand open competition) and other UK manufacturers of satellite components to meet the demand.

A government announcement on the subject is expected soon and it may seem rather bland, leaving the question of monopolies at least until the Home Office has reported on direct broadcasting by satellite. A decision will have to be taken fairly shortly, however, on the scale of Britain's effort in another space application — remote sensing — if it is not to miss the opportunity of cooperating in ESA's next programme. This year promises to be a vital one for Britain's space industry.

Judy Redfearn

Herbicide safety

Bill of health

The herbicide 2,4,5-T has been given a cautious but clean bill of health by two recently published reports. One*, by the Advisory Committee on Pesticides of the UK Ministry of Agriculture, Fisheries and Food, says that there is no sound medical or scientific evidence that herbicides based on 2,4,5-T are harmful to humans, animals or the environment in general. The second, by the European Community's Advisory Committee for Safety, Hygiene and Health Protection at Work, says there is no conclusive evidence that 2,4,5-T causes cancer, but asks for further evaluation of the long-term risks.

The British pesticides committee, essentially the licensing body for pesticides and herbicides, has reviewed 2,4,5-T nine times since 1970. Its latest review was undertaken at the request of the Minister of Agriculture after the National Union of Agricultural and Allied Workers claimed, in March 1980, that 2,4,5-T was harmful.

The union reviewed the scientific literature on 2,4,5-T and referred to 20 cases where it was alleged to have harmed humans or farm animals. The advisory committee says, however, that the union's evidence does not indicate that 2,4,5-T is a health risk.

Concern about 2,4,5-T has centred mainly on the presence of a contaminant, 2,3,7,8-tetrachlorodibenzodioxin (dioxin), a teratogen and carcinogen in some animal species. The committee now believes that this concern may have been misplaced and that the risks posed "by dioxin contamination in 2,4,5-T formulations may hitherto have been overestimated".

First, the committee says, dioxin contamination of 2,4,5-T formulations sold in the United Kingdom is now at the low level of 0.01 p.p.m. Second, the committee says that new studies enable it to identify for the first time "a daily level of intake below which effects on reproduction do not occur in the rat", that is 0.001 μg kg^{-1} day^{-1}.

In the circumstances, the committee considers that 2,4,5-T itself would present a problem before its dioxin contaminant, but that there is no convincing evidence that any effects caused by 2,4,5-T will be passed on to succeeding generations "on an heritable basis". The committee also notes that the WHO/FAO authorities have set a "no effect level" for 2,4,5-T in animals at 3 mg kg^{-1}. Employing a thousandfold safety margin, WHO/FAO have set a "temporary acceptable daily intake" for a man at 3 μg kg^{-1} for 2,4,5-T containing 0.1 p.p.m. dioxin.

The committee does, however, accept the union's claim that workers using 2,4,5-T are not always adequately protected. It suggests that exposure to the herbicide should in future be measured in urinary excretion.

On the question of alternatives to 2,4,5-T, the committee is doubtful. Much less is known about their toxicity and the committee, in continuing to allow 2,4,5-T to be used, is relying on the maxim "Better the devil you know . . .".

All of the twenty cases where 2,4,5-T exposure is alleged to have caused health problems are discussed in the report. The advisory committee has harsh things to say about coverage of alleged 2,4,5-T incidents and accuses the press of causing needless distress by publicizing cases without the consent of the individuals involved. The committee may have been unwise in making this accusation, given that most of the individuals referered to in the report seem to have sought out journalists.

Although the committee has given 2,4,5-T herbicides a clean bill of health, it does ask for prospective epidemiological studies of the risk from exposure to herbicides in general. Professor Robert Kilpatrick, chairman of the advisory committee and dean of the school of medicine at the University of Leicester,

---

**Further review of the safety for use in the UK of the herbicide 2,4,5-T.* Available free of charge from Pesticides Branch, Ministry of Agriculture, Fisheries and Food, Room 678, Great Westminster House, Horseferry Road, London SW1P 2AE, UK.

0028-0836/81/010004-02$01.00

says that a prospective study is in the offing and that 2,4,5-T would almost certainly be one of the herbicides investigated.

The report ducks one issue, however, that of trade union involvement in its decisions. The unions would prefer the Health and Safety Executive (HSE) — where the unions are included on various committees — to be responsible for assessing herbicide safety. Kilpatrick says that this is a matter for ministers to decide. But it seems likely that the unions will be involved in planning the prospective epidemiological study on herbicides. It remains to be seen, however, whether participation on this *ad hoc* basis will satisfy them. Despite the reassuring conclusions of the advisory committee's report, the recent decision by the Trades Union Congress to instruct its members not to handle any 2,4,5-T imported into Britain stands for the time being. **Alastair Hay**

UK forestry policy

Research wanted

British policy on forestry research came in for a drubbing just before Christmas in a critical report from the House of Lords Select Committee on Science and Technology. And the annual report of the Forestry Commission, the public agency responsible for publicly-owned forests, did little to answer the criticisms when it appeared a few days later.

The select committee's chief complaint is that there is no coordinated policy on forestry research. The Forestry Commission is accused of being too concerned with wood production to bother promoting less obvious areas of research, so that the Natural Environment Research Council has been left to support "fundamental and strategic research" — a responsibility which it has shouldered "lightly", according to the committee.

The committee asks that the Forestry Commission should mend its ways by appointing a chief scientist, and that the Advisory Board for the Research Councils should decide which of the two obvious candidates — the Natural Environment Research Council and the Agricultural Research Council — should be responsible for longer-term research. The commission should also be urged to provide advisory services for the owners of private woodlands much in the spirit in which the Agricultural Advisory Services help private farmers to tackle agricultural problems.

The commission's own account of the past year (to March 1980) confirms a preoccupation with the production of wood and in this respect some of the results are impressive. Machinery and pesticides have made the planting of trees a more certain business, with obvious benefits to production costs. The development of local seed supplies, the search for more productive species and for the avoidance of damage by high winds and the control of pests and forest diseases remain the mainstays of the commission's research.

Lords look ahead

The House of Lords Select Committee on Science and Technology will study scientific advice to government in its next inquiry. The idea came from Lord Todd, chairman of the committee. The committee will consider whether the post of chief scientific adviser to the government, which was abandoned in 1974, should be reinstated and whether responsibility for science should be removed from the Department of Education and Science.

British forestry is a controversial business with the commission frequently pilloried for its preoccupation with conifer plantations. The select committee makes a case for the preservation and renewal of often ancient broad-leaved woodlands, suggesting local authorities as suitable custodians. There is a plea for research to throw light on the more effective planning of land use.

The implementation of many of the select committee's recommendations will cost money — both for reorganization and for an increase in volume of some areas of research. The current level of forestry research funding of just under £5 million for the past year is not generous when compared with the current value of the forest estate of £500 million. But recognizing that an increase in public funds is unlikely, the committee suggests that extra revenue could be created by a 0.1 per cent levy on imported timber and wood products, raising £2.75 million a year.

There is little prospect that extra funds could be found in the Forestry Commission's profits. The commission acknowledges that over the next three years it is unlikely to make the financial target of 3 per cent a year in real terms required of it by the British government. In this respect, the Forestry Commission has yet to make an effective response to the criticisim of its financial operations by the Treasury in 1972. **Judy Redfearn**

UK clinical research

Future not so rosy

Celebrating their fiftieth anniversary in London last month, members of the Medical Research Society were not altogether optimistic about the future. One of the principal fears, aired at the day-long symposium on the organization of medical research, was the declining ability of clinical departments to attract and train the people to keep them in the front line of research.

Part of the problem is to provide the money and status to keep non-clinical research workers in clinical laboratories, where their knowledge and expertise is essential. Dr J. L. Gowans, secretary of the Medical Research Council (MRC) pointed out that although MRC grants can sometimes cover the costs of such people, the universities should really be providing them with permanent jobs.

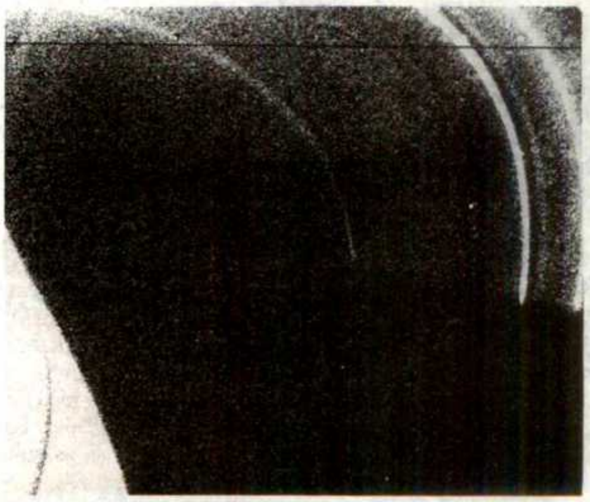
Professor W. S. Peart of St Mary's Hospital Medical School, London, said that in hard times, hard choices must be made in medical schools between technical staff and academic staff — it may be necessary to take on a new postdoctoral fellow rather than a lecturer.

The other side of the problem is the difficulty of providing high quality scientific training for medically qualified staff without disrupting the smooth course of their clinical careers. Professor Peart was disturbed by the lack of young people entering academic clinical research.

The continuing revolution in biology will increase the need to unite basic and applied research on man. Dr S. Brenner, director of the MRC Laboratory of Molecular Biology in Cambridge, pointed out that human molecular and cell biology, now in its infancy, will become the essential interface. In twenty-five years, he said, it will be as essential for researchers in pathology, cardiology and so on to have a thorough knowledge of cell and molecular biology as it is now for clinical endocrinologists to know their steroid chemistry. **Mary Lindley**

New rings for old

Claims from ground-based astronomers that Saturn has an innermost D ring, stretching from the bright C ring down to the planet's cloud tops, were apparently disproved by Pioneer 11 in 1979. But this image from Voyager 1, taken soon after its encounter with Saturn last November, reveals an inner ring, complete with fine structure, between the planet (bottom left) and the C ring (top right). Here, Saturn's shadow falls across the ring, which is apparently too faint to be detectable from Earth.



0028-0836/81/010005-01$01.00

UK universities

Feeling the pinch

Figures released by the Committee of Vice-Chancellors and Principals show that the number of overseas undergraduate and postgraduate students admitted to British universities in October 1980 was 90.7 per cent and 88.9 per cent respectively of the number admitted in October 1979. The intake of British undergraduates was 2.3 per cent higher than in 1979, an increase in line with the rise in the number of 18-year-olds in the population.

The shortfall in the number of overseas students was not as great as some had expected after the British government's decision earlier in the year to charge them the full economic fee for their education, averaging about £3,000 per year. But recent figures released by the Universities Central Council on Admissions (UCCA) indicate that a much greater shortfall may be expected this year. The number of applicants from overseas for admission in October 1981 is 33 per cent down on the number received at the same stage last year. One reason for the delayed reaction may be that governments had already committed themselves to paying for students to study in Britain this year before the increased charges were announced. Applications from British students for university places in 1981, however, continue to show a small increase, the number from men being up by 2 per cent and from women by 7 per cent.

Figures released at such an early stage in the academic year are not a reliable guide of what will happen next year. Nevertheless, many universities are worried. They will face a difficult decision in February and March when they have to fix their fees for 1981–82. With a considerably reduced income from overseas students, they will have to decide whether to compensate by increasing the fees of overseas students yet again at the risk of discouraging even more students from accepting places.

According to the vice-chancellors' committee, the increase in the number of home students only adds to the universities' problems because the government has set aside no extra funds to meet the growth. Well qualified home students could well be refused places because of a lack of cash to support them. The full effects of the increase in overseas students' fees will be felt in 1982–83 by which time government protection of overseas students who began their studies before October 1980 will have ceased. The government will then have effectively removed 10 per cent of universities' income from the recurrent grant.

The Committee of Directors of Polytechnics reports a similar but more marked trend for entry to full-time and sandwich courses in October 1980. Admissions of overseas students were down 29 per cent on 1979, but this was more than compensated for by a 7 per cent increase in home students, much of the growth being in science and technology courses.

Some polytechnics with a high percentage of overseas students, especially those near London, may be hit particularly badly. Others may only be able to offer those courses for which there is a buoyant demand. At Portsmouth Polytechnic, for example, the masters' course in fuel technology, the only one of its kind offered in Britain, is very popular. But courses offered by several institutions may suffer badly from increased competition.

polytechnics' advantage may, however, be that they can adapt their courses to changing demands more easily than the universities.

Few universities and polytechnics have yet analysed this year's intake in terms of subject area and students' nationalities. But general impressions are that students from South-East Asia are having more difficulty in raising fees than those from the Middle East and that science and engineering courses are suffering less than courses in the humanities.

Judy Redfearn

Poland freezes Antarctic research

Poland's Antarctic research programme for 1981 will include a "wide range of economy measures" according to a statement last month by Dr Maciej Zalewski, Secretary of the Polar Research Committee of the Polish Academy of Sciences. From now on, he said, research will concentrate on "strategic lines" of special importance to science and the economy.

The main saving, said Dr Zalewski, will be made by discontinuing large and expensive summer expeditions, and concentrating on small teams occupying a permanent base on a yearly shift system. Nevertheless, some summer activity will continue — such as the current joint expedition from the Academy of Sciences and the Marine Fisheries' Institute in Gdynia which is part of the international "Biomass" programme, as is the similar expedition planned for summer 1982–83.

Arctowski Station, King George Island.

What, then, is to be cut? According to Dr Stanislaw Rakusa-Suszczewski, of the academy's Institute of Ecology, very little of any significance. Antarctic programmes are planned by the three relevant institutes of the academy — geophysics, biology and ecology — and financed by the academy, and judging the proposed research entirely on its intrinsic merits, the arguments were so strong, he said, that the academy has decided to go ahead. Under the new programme, the main emphasis will be on marine biology and the study of whales, but in all disciplines there would be increased stress on fund a mental research.

This represents a significant change in research planning. Since the early 1970s, research in Poland has been target-oriented and financed on a project basis, within a complex hierarchy of "problems" of the national economy. (Fundamental research was still possible, but had to be formally associated with some specific "problem"). The new intellectual climate in Poland, with its emphasis on greater academic autonomy, described by Dr Rakusa-Suszczewski as a "very positive change", has clearly meant that, in reshaping the Antarctic programme to fit Poland's reduced economic circumstances, the planners of the academy have been able to make their decisions on basis of scientific value.

Vera Rich

Brazil's space programme

Aiming high

Saõ Paulo, December

Showing refreshing optimism at a time of unprecedented inflation, plans have been drawn up in Brazil for an ambitious space programme costing hundreds of millions of dollars over the next ten years. Remote sensing and meteorological satellites and the rockets to launch them are included in a plan approved in principle by the Brazilian government.

Meteorology and remote sensing have already attracted substantial interest and investment in Brazil, where the information they can provide on the physical conditions of little-known areas is especially valuable. The remote sensing programme, set up in the 1960s, was intended to help map the Amazon Basin. The programme has grown from the use of aerial survey to using Skylab and Landsat images. A Landsat receiving station was

0028-0836/81/010006-02$01.00

opened in Curitiba in 1975 and now provides 20,000 images a year to 1,100 users, many of them in other South American countries. The next step will be to upgrade the station, at a cost of about $6 million, to receive Landsat D images.

The meteorology programme involves the building of up to 100 data-collection platforms around the country. A prototype is being tried out in France and the first Brazilian platforms are expected to start operating early this year. Information from them will be relayed by the American TIROS N and GOES satellites. A major objective of the platforms is to provide accurate forecasts of drought in the impoverished north-east region.

According to Dr Jesus Parada, director of the civilian space research institute, INPE, previous cooperative ventures, mainly with America, have given Brazil the necessary expertise to go it alone.

The intention is to launch four satellites a year from 1986–87. INPE would develop the satellites and the launcher would be developed at the nearby military space institute. Total cost of the programme is estimated at more than $700 million, about two thirds of it for launcher development.

The first two satellites would take over from TIROS N and GOES by relaying information gathered by the meteorological data collection platforms to a national centre. And the number of platforms could be increased to several thousands if the new satellites are launched. A radiation budget experiment is also planned. The first satellite, weighing 150–200 kg, would be launched on the qualification flight of the rocket into an orbit of 25–30° at an altitude of 700 km.

The third and fourth satellites would be more complex, devoted to direct land imaging and weighing 250–300 kg. They would be placed in a near-polar orbit at an altitude of 650 km.

All four satellites would be launched from a site in Brazil, possibly the existing military launch pad near Natal. A major drawback of this site, however, is that large payloads would have to be launched into an unusual trajectory to minimize the threat to the city which is only a few kilometres away.

Ambitious though the new plans are, progress so far has been slow. With a budget of only $10 million for 1980, feasibility studies have not got as far as Dr Parada would have liked. But if the hoped-for budget of $40 million is forthcoming for 1981 the project can move on into the design phase.

In a country with grave social problems there are many areas where investment might provide a more obvious and immediate return than a venture into space. And a second space project, for Brazil's first telecommunications satellite, is also competing for funds. So the chances of Brazil's space programme are closely tied to the fate of Brazil's economy over the next few years.

Judy Redfearn

Nuclear waste

US Senate stalls

Washington

Hopes for new legislation establishing a national policy for the disposal of radioactive waste — eagerly sought by both the nuclear industry and the environmental movement — foundered in the Senate last month on disagreement over whether states should be involved in decisions about military waste from nuclear weapons construction.

Main responsibility seems to rest with Senator Henry Jackson, powerful chairman of the Senate Energy committee and a member of its Armed Services Committee.

Senator Jackson had previously supported a proposed nuclear waste bill that was passed in the Senate in July. This bill contained limited provisions for the involvement of states in siting decisions, which the Senator opposes, but it was sweetened by the inclusion of plans for "away-from-reactor" sites which the owners of nuclear reactors could use for the temporary relief of their own spent-fuel storage tanks.

The House of Representatives, after lengthy negotiations between pro-nuclear and anti-nuclear forces, passed a compromise bill last month. Like the Senate bill, it provided for a state to veto a siting decision, if it obtained support of one of the two congressional bodies.

However, the House bill did not include any mention of away-from-reactor storage, which the industry would like since it would keep spent fuel available for future reprocessing. And Senator Jackson made it clear that he did not support the bill, killing any chance of agreement by offering an amended version unacceptable to the House.

The storage of nuclear waste remains the Achilles' heel of the nuclear power industry, with several states currently refusing permission for further power-plant construction until adequate means for long-term storage are available.

The nuclear industry itself is keen to find more permanent storage for the 80,000 metric tons of spent fuel now lying in storage tanks next to reactors around the country. And the environmental movement wants a nuclear waste policy that is both ecologically acceptable and open as much as possible to local participation in decision-making.

In February, President Carter announced plans for a long-term waste disposal policy. Reiterating his previous opposition to reprocessing, this is based on the permanent storage of waste in deep geological formations — a technique considered both feasible and relatively safe by most of the scientific community.

Carter proposed that long-term studies of possible sites should be initiated, with the goal of having a permanent disposal facility in operation by 1995. But the Senate, acting under pressure from the nuclear industry, expressed impatience with this schedule, and the bill which it passed in July emphasized short-term waste storage procedures.

Sharp differences also occurred in the House of Representatives. Pro-nuclear forces, led by Congressman Mike McCormack of the Science and Technology Committee, supported a bill which would require the Department of Energy to construct technology demonstration facilities in various parts of the country, to serve as a basis for licensing a full-scale commercial waste facility.

In contrast, the chairman of the House Interior Committee, Congressman Morris Udall, supported a bill more closely modelled on the Administration's proposals, arguing that it was premature to rush into demonstration projects.

In the end a compromise was reached between the two committees and the House passed a joint bill, under which the Department of Energy would pick two possible burial sites by 1982 and two more by 1985; and by 1987, following an elaborate series of public hearings, the President would recommend to Congress the site or sites which he considered safe for the storage of waste.

Despite the faster schedule, this compromise proved unacceptable to the Senate. Stung by the omission of any provision for short-term away-from-reactor storage, Senator Jackson told the Senate that without a specific exemption for atomic defence repositories, the House bill as it stood infringed the jurisdiction of both House and Senate Armed Services Committees, and could conflict with existing clauses in the Atomic Energy Act.

Environmentalist groups are insistent that military wastes, amounting to up to 90 per cent of the nation's high-level radioactive waste, should also be covered in any national plan.

But the Senate's decision now means that the process of developing legislation must start over again when the new Congress convenes this month. In the Senate at least, recently-elected conservative legislators are likely to be more sympathetic to the nuclear industry.

Meanwhile, in line with the Carter Administration's long-term waste disposal policy, the Department of Energy has issued an environmental impact statement confirming its view that the best method of permanently disposing of radioactive wastes from nuclear power plants is to place them in mined depositories deep in geological formations.

According to the department, a typical repository would require about 2,000 acres underground, and above-ground facilities would occupy 500 to 750 acres. Access roads would take up another 30 acres. A total of 2,000 acres above ground would be needed for each facility, with consequent restrictions of mineral and surface rights.

David Dickson

0028-0836/81/010007-01$01.00

CORRESPONDENCE

Badgers and TB

SIR — I refer to the letter by Dr Stephen Harris in your issue of 11 December 1980 (p.532) which, while critical of certain aspects of Lord Zuckerman's Report to the Minister of Agriculture on "Badgers, Cattle and Tuberculosis", agreed with his basic conclusion that the badger is a major reservoir of bovine tuberculosis (TB) in some areas of South-West England. The main criticism was that, while the reservoir constitutes a potential danger to cattle in limited areas, there is no "scientific evidence" to justify the deduction that the level of tuberculous infection in badgers in the South West is "dangerously high" and may result in spread of the disease to adjacent populations.

Since Dr Harris writes for and on behalf of the Mammal Society, it may be appropriate to ask whether all the interested members of that society would fail to be convinced by the evidence. More than 4,000 badgers were examined in connection with official investigations in the years 1971–79 and 14 per cent at least harboured *Mycobacterium bovis* (Report: para. 130). Of animals found dead since 1972 in fields, woods and farms in Gloucestershire, Avon and Wiltshire, 54/194 (28 per cent) were found to be tuberculous. Of those killed in Gloucestershire, predominantly on the roads and presumably by chance, during this period 1976–80, 21/232 (9 per cent) were tuberculous; this admittedly was the highest recorded prevalence by county but for the whole of the South West the figure was still 4.2 per cent (Report; Table 3).

These figures should be a cause for concern when dealing with a social animal which occupies, for much of the time, confined spaces underground and whose movements above ground favour between-group contacts. Bovine tuberculosis in badgers is sometimes a rapidly progressive, lethal disease, the causal organism probably being excreted by many routes in large numbers[1,2]. It has, therefore, ample opportunities for transfer and by any objective epidemiological standard the risk of spread to neighbouring badger populations must be considerable. The onus for showing that this is not a reasonable interpretation should rest on others.

The argument advanced by Dr Harris, that the recent decline in the incidence of tuberculosis in cattle in the South West is simply a reflection of a general downward trend in the figures for England as a whole, misses the point that the incidence of herd "breakdowns" in the South West, in the years 1974–80, has never been less than five times that for the rest of England (Report; Table 19).

Many anomalies in the data were highlighted in the Report, which pointed out the need for continuing investigation and research. In this connection it is difficult to see how the "subtle factors" twice mentioned by Dr Harris as significant in the spread of the disease from infected badgers, could be elucidated without extensive sampling on a statistically significant scale. In particular, continuous monitoring of badgers on the perimeter of heavily infected areas would be necessary to prove that the disease is *not* spreading (Report; para. 137). It is to be hoped that consideration will be given by interested bodies as to how this could be organized, if the necessary funds were provided.

From the Mammal Society:
In the letter from the Mammal Society on badgers and TB published on 11 December, the statements appear to be attributed to Stephen Harris. I wish to make it clear that the letter is a Mammal Society view and that Stephen Harris was only personally involved in the communication of the letter to the Editor.
— J. R. Flowerdew, Hon. Secretary.

W. PLOWRIGHT

Department of Microbiology,
ARC Institute for
Research on Animal Diseases,
Compton, Berkshire, UK

1. Gallagher, J., Muirhead, R.H. & Burn, K.J. *Vet. Rec.* **98**, 9-14 (1976).
2. Gallagher, J. & Nelson, J. *Vet. Rec.* **105**, 546-551 (1979).

Museum policy

SIR—The broader implications of cladism are not my special concern, but the study of human evolution is. Cladistics is a useful research tool, but I support L.B. Halstead's[1] objections to the way it is used in the new exhibit *Man's place in evolution* at the Natural History Museum.

Cladistic analysis is still a relatively untried method for expressing relationships between groups, and technical details of its application are still being debated. It was devised as a conceptual framework within which one aspect of the relationships between modern and fossil forms could be analysed. It relates groups solely on the basis of shared features, and rejects the notion that any ancestral relationship can be inferred from this evidence; such relationships are considered "unknowable". Proponents claim that only by adhering to these principles can palaeontologists frame testable hypotheses. The strengths of the method are its simplistic rigour, but therein also lie its weaknesses. In order to provide an unambiguous framework, certain assumptions have to be made. For example, speciation is recognized only as a dichotomous branching event, and more complex adaptive radiations, anagenesis and convergent evolution of morphological features are discounted. Such assumptions are unrealistic. The concept of "relatedness" enshrined in cladistics is therefore fundamentally different from the general one which implies a lineal relationship, and as used in the exhibit, it is likely to confuse, if not actually mislead, the visitors.

The contents of the exhibit also give cause for disquiet. The authors seem particularly confused about what constitutes the "habiline" and "australopithecine" groups of early hominids. Two important specimens "1470" and KNM-ER 1813, have been seriously misassigned. KNM-ER 1813 is consistently cited as an example of a "habiline" and "1470" as an "australopithecine". Yet, "1470" has a cranial capacity of around 775 cm^3 and is included in *Homo habilis* by many workers. The cranium KNM-ER 1813 is smaller overall, and it has a cranial capacity of between 500 and 550 cm^3. It has not yet been formally assigned to a taxon but its position is sufficiently enigmatic for it to have been compared, by some authors, to *Australopithecus africanus*, and by others, to early *Homo*. Howell[2] has attributed KNM-ER 1813 to *Homo habilis*, but he also includes "1470" in the same species.

The section which deals with *Homo erectus* also deserves comment. Recent, and well known, discoveries in Europe and Africa have demonstrated that a much wider range of cranial morphological features is subsumed into the "erectus" group than was previously believed. Several skulls shown an apparent mixture of classic "erectus" and archaic *Homo sapiens* features. These new finds strongly suggest a continuous gradation of morphology from "erectus" to "sapiens". The exhibit totally ignores this evidence.

To attribute important finds wrongly is careless; to ignore "uncomfortable" evidence is dubious academic practice, and to imply that cladism is an orthodox systematic approach is irresponsible. The contribution of the Natural History Museum to science and education is too important to be compromised in this way.

B.A. WOOD

Middlesex Hospital Medical School,
London W1, UK

1. Halstead, L.B. *Nature* **288**, 208 (1980).
2. Howell, F.C. in *Evolution of African Mammals* (eds Maglio, V.J. & Cooke, H.B.S.) 154-248 (Harvard University Press, Cambridge, 1978).

SIR — In a recent letter to *Nature* (**288**, 208; 1980) L. B. Halstead states his convictions that human evolution was gradual, that its gene pool of the past had certain knowable characteristics, and that the ancestry of a living species can be determined with deadly accuracy. Halstead's convictions arise from his discovery of a new form of *doubt*, such that he can contend that "there is not any serious doubts about *Homo erectus* being directly ancestral to *Homo sapiens*". This is certainly the news Biology has waited for, the moment when the Truth can at last be known so that all this difficult and extremely tiresome theory can be dispensed with. Until now my colleagues and I had always imagined that to doubt something was "to be uncertain as to a truth or fact" and the notion that distinguishes science from, say, politics is that in science uncertainty about the truth must remain or progress ends. Halstead's discovery can only mean that he has in hand a new form of truth — a kind of truth that can be known. Many of us puzzled over what kind it might be until it finally dawned on us that it emerges from false doubt and, to honour its discoverer, I call it *Halsteadian Truth*. If ever again we are troubled by some theoretical matter we need only consult the oracle in Reading. Now that Halsteadian Truth permits us to know at last the way in which evolution proceeds, what extinct gene pools were like, and exactly who was whose ancestor, I confess, with sadness for years misspent, that cladistics is indeed a waste of time.

This news about the emptiness of cladistics will have a profound effect on systematics, for

Continued on p. 105

The ocean floor closely observed: will it subvert marine geology?

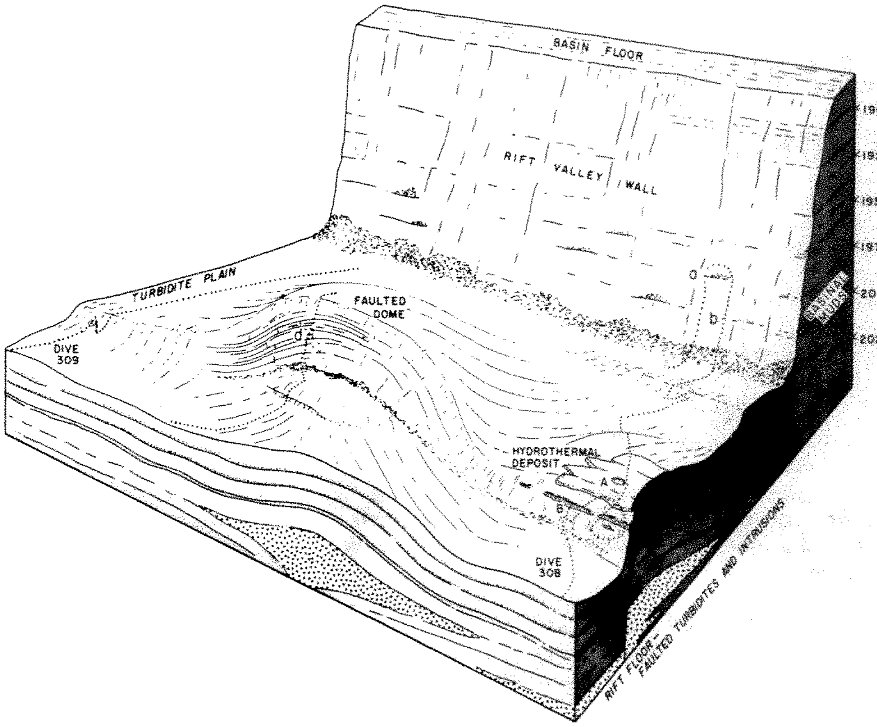
from Tjeerd H. van Andel

In a recent paper based on work with a deep-diving research submarine, fine-scale observations of the sea floor are ingeniously extrapolated to models of the events occuring at depth*. The paper is a good example of the geologist's method of using detailed observations on rock outcrops to yield grand interpretations of causes, and it is as such quite similar to the usual fare in geological journals except for an important difference — the observations were made 2,000 m below the sea surface rather than on a walk in the hills.

For a century, marine geologists have coped with severe limitations imposed by the vastness of their domain, by the filter applied to their observations by water depths of several kilometres, and by the difficulty of knowing where one is. The first of the constraints produced descriptions of entire ocean basins equivalent to a geological synthesis of Britain based on ten outcrops, while the second and third limited resolution to features tens or even hundreds of metres across. The traditional oceanic mapping scale is about 1:1,000,000, and in the 1960s a 'detailed' map was usually drawn to a scale of 1:250,000. We did, of course photograph the sea floor but the gap between a photo of a few square metres and a map representing the terrain at a scale of 2.5 km to the centimetre is near unbridgeable. We observed mid-ocean ridges and sand ripples but very little in between. Until a decade ago, the divergent plate boundary appeared as a V-shaped cleft on sounding records; now we know that it consists of vertical fault scarps, terraces, flat lava floors, rows of small but complex volcanoes and fascinating hot springs.

About a decade ago our ability to appreciate detail improved rather suddenly with the introduction of instrument vehicles towed near the ocean floor, manned submarines and, above all, with the means to determine positions within ten metres or so. Since 1970, large expeditions to several mid-ocean ridges have demonstrated the power of the new technology with numerous startling discoveries, and have introduced new ways of thinking about marine geology that are akin to traditional approaches in petrology and structural and sedimentary geology on land. At times the change has been disconcerting to marine geologists not accustomed to looking at their world on a fine scale, but the new opportunities clearly hold much promise.

The impact of this costly and logistically complex technology has not been restricted to scientific data and inferences. There has been some sense of discouragement. A decade of hard work has covered little territory at a high cost. Perhaps 0.01 per cent of the divergent plate boundary has now been studied in detail at a cost of much more than 10 million dollars. One wonder whether, at this rate, we shall ever be able to obtain a reasonable sample of the rest of the ocean floor. The work is also



A sketch of the undersea terrain near to the north west wall of Northern Trough, Guayanas Basin, Gulf of California from two dives made by Lonsdale and Lawver in the submersible *Sea Cliff.* The Gulf lies between two crustal plates that are drifting apart, in a position like that of the Mid-Atlantic Ridge between the American and the European and African plates. At mid-ocean ridges, as the plates separate, molten rock wells up from below to fill the rift and heal it, forming a string of small volcanoes. In the Gulf of California the same process occurs but is modified by a thick blanket of sediments shed from the nearby Continental margins. As a result the lava does not reach the surface but forms mushroom shaped intrusions within the sediment layer and deforms and domes the overlying deposits and sea floor.

---

*Lonsdale, Peter, and Lawver, L.A. *Immature plate boundary zones studied with a submersible in the Gulf of California; Geol. Soc. America Bull.*, **91**, I, 555 (1980).

Tjeerd H. van Andel is Professor of Oceanography in the Geology and Geophysics Departments of Stanford University, California.

0028-0836/81/010009-02$01.00

cumbersome and improving the instrumentation is sometimes more challenging than using it. This is not without reason. The paper that provoked this discussion, based on minimal technology, suffers the weakness of a consequently much reduced precision.

Coupled with this image of an overwhelming task is a feeling of insecurity. Does any value, except crude geographic description, remain in the work done before? Are the sparse observational nets, the data of low resolution still useful? Does the image of the ocean floor obtained over the last century still hold up? Of course no one was ever wholly unaware of the shortcomings of our methods, but we coped with that by assuming that, somehow, the ocean floor was simpler than the land, the gradients of change more gentle, the processes not so variable in time and space. We thought subconsciously of the sea floor as having the gentle and smooth slopes of our sounding records and were startled by the discovery of vertical fault scarps and of open fissures. Similarly,the petrologists, having attributed to variations in the earth' mantle the subtle differences between basalts from various parts of the mid-ocean ridge, were surprised to find this entire range exhibited in a single 3 km long traverse across the volcanic axis, presumably as a result of ageing of the magma chamber. In a limited way, but we have no other means, we can test the value of the old data by inquiring after their impact on our present understanding of the geology of the mid-ocean rift. I submit that this impact is negligible and no reason for comfort.

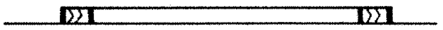
It does not seem proper to retreat one step and say "Well, the mid-ocean ridge is an active and hence complicated place but, surely, elsewhere in the ocean things are much more uniform and do not need this sort of detail". This, then, suggests that we should ask whether the detailed, fine-scale, near-bottom approach is the way of the future, rendering obsolete most of the old ways that bemuse our colleagues on land but are second nature to us. Aware that I am putting it very strongly I propose that the answer is yes and that we should, the sooner the better, divert most of our resources to this approach and leave the large scale global mapping to organizations charged with mapping. I am willing to allow for numerous exceptions but it seems that we have gathered enough regional data now and must proceed with the methods that provide the best hope to answer fundamental questions of process and history. I do not truly believe that we shall soon see the last of the roaming marine geologist disappear but I think that we should attempt to hasten his demise in order to look at the ocean floor in the detailed, time-consuming, expensive and, alas, often tedious way that will give us the answers we need. □

Did retroviruses evolve from transposable elements?

from Andrew Flavell

RETROVIRUSES (formerly called RNA tumour viruses) have attracted much interest as they can cause cancer and are ubiquitous in a wide variety of vertebrates. It has long been known that DNA copies of these RNA-containing viruses are synthesized in infected cells and these copies are able to integrate into the cellular DNA, becoming part of the genetic complement of the cell. Recently the structures of the integrated 'provirus' DNA copies of several retroviruses have been studied by a combination of DNA cloning and sequencing methods (Shimotohno *et al. Nature* **285,** 550, 1980; Dhar *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 3937; 1980; Sutcliffe *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 3302; 1980; Majors *et al. Cold Spring Harb. Symp. quant. Biol.* **45,** 1980; Benz *et al. Nature* **288,** 665; 1980; Czernilofsky *et al. Nucleic Acids Res.* **8,** 2967; 1980). These results have shown that there exists a striking similarity between these structures and those of 'transposable genetic elements' from both prokaryotes and eukaryotes. Transposable elements themselves are a highly interesting class of sequences (see *Nature* **287,** 188; 1980). In bacteria these sequences are able to move around inside the host cell genome and can catalyse chromosomal deletions and rearrangements. In *Drosophila* they can probably act as insertional mutagens (Gehring and Paro *Cell* **19,** 897; 1980; Rubin and Bingham, personal communication). Furthermore, the generation of antibody diversity in mammals, the control of mating type in yeast and alteration in surface antigens in trypanosomes all occur via DNA rearrangements which may share features in common with those of transposable elements. Lastly, tumour progression is often accompanied by chromosomal rearrangements, though no evidence exists connecting such changes with transposable elements; indeed their presence has not yet been proven in vertebrates.

All eukaryotic transposable elements and retrovirus proviruses studied so far possess DNA sequence repeats a few hundred nucleotide pairs long at each end of the element (see the figure). This structure is also shared by some bacterial transposons although there seems to be a much broader spectrum of transposable structures in bacteria. It therefore seems tempting to postulate a familial relationship between all these various species. This idea has been leant further weight by DNA sequencing of several eukaryotic transposable elements and integrated retrovirus proviruses at the presumed key regions for transpositon, namely the ends of the elements. Here, the (so far) inviolable rule is the presence of the dinucleotides TG . . . . . CA at the ends of the element. This homology is much greater in several cases, notably that of the *Drosophila* element *copia* which shares with an avian retrovirus the sequence TGT . . . . . TACAAACA at it's ends (Levis *et al. Cell* **21,** 581, 1980; Shimotohno *et al. Nature* **285,** 550, 1980). Such similarities are very unlikely to be coincidental but cannot be regarded as proof of an interrelationship between these far-diverged species. Not surprisingly, some effort is being expended to relate these two fields by discovering whether eukaryotic transposable elements are really a special form of integrated retrovirus provirus or whether retroviruses are, in fact, transposable elements with an extra trick up their sleeves. The article in this week's *Nature* by Keshet and Shaul (see page 83) describes a family of DNA sequences in mice that may be both types of element simultaneously. Originally described by Howk *et al.* (*J. Virol* **25,** 115; 1978) these sequences (termed VL30 genes) were assumed to be retrovirus-related (although no sequence homology was detected with any retrovirus) by the ability of their RNA transcripts to be packaged into helper retrovirus virions which could be used to infect other cell types. Keshet and Shaul have now shown that those sequences possess direct DNA repeats at their ends. What is the relationship of this sequence family to the proviruses and transposable elements?



The generalized structure of a eukaryotic transposable element. The repeat units (>) are in the same orientation as the DNA and each repeat is itself bounded by small inverted repeats (|) of between 2 and 12 nucleotide pairs.

In 1970 Temin discussed in his protoviral hypothesis the possibility that retroviruses might have evolved from an ancestral gene by successive cycles of transcription, reverse transcription and reintegration (*Persp. Biol. Med.* **14,** 11; 1970), each cycle picking up extra genetic information. In this light, such sequences as the VL30 family and perhaps some defective retroviruses might be regarded as intermediates in this pathway. It also seems plausible that retrovirus evolution might have proceeded via DNA transposition,

---

Andrew Flavell is in the Imperial Cancer Research Fund Laboratories, Mill Hill, London.

0028-0836/81/010010-02$01.00

each new transpositional event adding to the repertoire of the element until its RNA transcript possessed the capsid and reverse transcriptase genes necessary for a successful sortie outside its home cell. Or possibly a combination of the two processes might be used. However, the answer to such crucial questions as 'Do RNA copies of the yeast and *Drosophila* elements ever leave their cells?', and 'Can retroviral proviruses move around their host genome without recourse to an RNA intermediate?' are as yet unknown. What is not in doubt, however, is that both types of species, by their ubiquity, sheer quantity (it has been estimated that 5 per cent of the *Drosophila* genome is composed of *copia*-like sequences) and intriguing properties are well worthy of further study. To prove that retroviruses and transposable elements can interconvert in either direction it will be important (and difficult!) to find examples of the 'missing links' between them, one of which may well be the V130 gene family described in this issue. □

In the beginning

from E G. Nisbet and C. T. Pillinger

Our exploration of Darwin's "warm little pond" in search of the origins of life on Earth has revealed but a few islands of fact. New and apparently firm ground slowly appears: sometimes it turns to quicksand. In this issue of *Nature*, Bridgwater *et al.* (see page 51) cast doubt upon earlier reports of evidence for life in 3,800-Myr old rocks from Isua, Greenland. Here we examine some of the new evidence concerning the existence of life in the early Precambrian, and discuss recent advances in the understanding of the nature of the early surface of the Earth.

When did life begin? In an earlier review, Nisbet (*Nature* **284**, 395; 1980) concluded that the oldest possible evidence of life was from Isua, the oldest probable evidence from 3,500-Myr old rocks in W. Australia and the oldest definite evidence from 2,700-Myr old strata. Much work has been done on the Isua rocks and three lines of evidence have emerged: Pflug and Jaeschke-Boyer (*Nature* **280**, 483; 1979) have described apparent 'microfossils'; hydrocarbons have been reported among chemical extracts (*Chemical Eng. News* **75**, 22; 1979); and isotopic studies on carbonates (Schidlowski *et al. Geochim. Cosmochim. Acta* **43**, 189; 1979) suggest biological controls on the carbon cycle.

Pflug and Jaeschke-Boyer detected cell-like inclusions in cherty layers from Isua. They investigated these by the innovative use of the Raman laser molecular microprobe, an excellent approach unifying the efforts of palaeontologists and organic geochemists. The microprobe showed spectral features indicative of aliphatic side chains and esters. From this and their petrographic study Pflug and Jaeschke-Boyer concluded that the inclusions were microfossils of a primitive organism which they termed *Isuasphera*. If so, this would be the oldest evidence of life.

Bridgwater *et al.* have examined duplicate samples of the same rocks. From morphological and petrographic studies and the known metamorphic history of the rocks, they conclude that *Isuasphera* are no more than fluid inclusions of secondary origin.

Much of the debate hinges on the metamorphism; Isua rocks may have reached temperatures as high as 500-600°C. Perhaps more important, oxygen isotopic data (Perry *et al. J. Geol.* **86**, 223; 1978) suggest minimum alteration temperatures of 390°C. Bridgwater *et al.* point out that high grade metamorphism does not entirely preclude the survival of microfossils. However, individual geolipids and esters could not be expected to remain intact after experiencing such conditions. In another paper querying Isua studies, Nagy *et al.* (this issue of *Nature* page 53) indicate the life expectancy of a straight chain aliphatic hydrocarbon at 600°C as about 10 seconds! In the oil industry it is well known that by 250°C everything that can be cracked will be cracked to give methane and graphite end products.

What then is the origin of Isua organics? Ponnamperuma (personal communication to C.T. Pillinger) has argued that hydrocarbons present in extracts of the acid resistant residues of Isua rocks must be contemporary because repeated and prolonged extraction of the intact rock matrix failed to release them (see also Barlow *Open Earth* **9**, 16; 1980 for a similar report). On the other hand Nagy *et al.* suggest that the extractables are no more than a few tens of thousands of years old and derive from encrusting lichens and cyanobacteria.

The lessons of contamination have been painfully learned by organic geochemists. Metamorphism may not be a problem of similar magnitude but it is noteworthy that Hayes *et al.* (Hayes, Kaplan and Wedeking, in preparation), from a painstaking study of H/C ratios in over 400 Precambrian samples, have concluded that all early Archaean samples investigated by them have been affected by thermal events. The H/C data suggest that if evidence of biological activity survives in early Archaean rocks, it must be within the kerogen (the cross-linked organic polymer) rather than extractable geolipids. Stable isotope studies by the same investigators on 50 samples, selected as worthy of kerogen isolation, should produce interesting results.

Thus recent work has thrown much doubt on the supposed evidence for life in Isua time. Nevertheless, there is carbonaceous matter, albeit graphite, in Isua rocks, and the third category of evidence, isotopic studies on the carbonates, remains intriguing. Schidlowski *et al.*'s work is not inconsistent with biological processes in a stabilised carbon cycle. The evidence remains 'possible'.

In slightly younger rocks, the information is becoming much more certain. There are three well-studied 3,500-Myr old terrains of low metamorphic grade — the Pilbara Block, Australia; the Barberton Mountain Land, South Africa; and parts of the Rhodesian Craton, Zimbabwe. In all three, a variety of sedimentary rocks exist, some of which are remarkably little deformed and metamorphosed, in contrast to the samples from Isua. In the Pilbara, new mapping, both by Hickman and Lipple (Geological Survey of W. Australia), and in the very important field program of the University of W. Australia, has led to the discovery of much evidence of 3,500-Myr life, especially in rocks collected by J. Dunlop and R. Buick (University of W. Australia). Dunlop *et al.* (*Nature* **274**, 676; 1978) found possible microfossils and Walter *et al.* (*Nature* **284**, 443; 1980) found a probable stromatolite structure. Significantly, important new discoveries have been made by Awramik, Walter, Schopf and co-workers in rocks from areas studied by Dunlop and Buick (*Geotimes* Sept, 1980). This work, "finding the first conceit of love there bred, where time and outward form would show it dead", is becoming accepted as sound evidence for early life. It may be that as early as 3,500-Myr ago there was an abundant, simple but diverse biological community in the seas.

So much for the islands of 'fact'. What of the warm pond of speculation? Recent modelling (often by Monte-Carlo methods) of early Archaean conditions is helping to place constraints on that pond. Frey (*Precamb. Res.* **10**, 195; 1980) and Grieve (*Precamb. Res.* **10**, 217; 1980) have suggested, though with different models, that late heavy bombardment of the Earth by asteroid sized bodies established a topographic dichotomy between 'continents' and 'oceans' by about 4,000-Myr ago. Former models proposing a global ocean, such as that of Hargraves (*Science* **193**, 363; 1976) appear unlikely; the process described in *Psalm* 104, 5-9 appears to have taken place very early indeed.

Models of the atmosphere have also

E. G. Nisbet and C. T. Pillinger are in the Department of Earth Sciences, University of Cambridge.

0028-0836/81/010011-02$01.00

improved. In recent years there has been a strong shift away from the traditional idea that the early atmosphere of the Earth consisted of a highly reduced mixture of methane, ammonia and hydrogen. If a reducing atmosphere did once exist, it would probably not have survived for long, ammonia being very rapidly converted by photolytic destruction to nitrogen. Hart (*Icarus* **33**, 23; 1978), Owen *et al.* (*Nature* **277**, 640; 1979) and Henderson-Sellers and Schwartz (*Nature* **287**, 526; 1980) have shown that the 'window of life' must have been rather narrow: if the atmosphere had had too much CO$_2$ or if temperatures had been slightly higher a Venus-like runaway greenhouse inferno would have developed; too little and Martian-like glaciation would have occurred. Perhaps only in a relatively short period around 4,000-Myr ago were surface conditions suitable for life to appear: once it had appeared, the CO$_2$ levels would be controlled and the danger of runaway averted. Henderson-Sellers and Schwartz also highlighted the ammonia 'problem' — there was probably very little in the atmosphere, except perhaps small quantities produced abiotically by photochemical reduction of N$_2$ catalysed by TiO$_2$ in desert sands. Consequently, those chemical evolutionaries who rely on the availability of ammonia may depend on the uncertain intervention of a *deus ex titania.* Life did appear, the warm pond was occupied, but it may have been a close-run thing. Perhaps the Earth is indeed a very rare place, perhaps there are few Gardens in Eden. □

Magic bullets

from Tony Davies

THE possibility of constructing Zauberkugeln — magic bullets — was first thought of by Paul Ehrlich in his inspired attempts to find new cures for a variety of human diseases. Magic bullets, he said, would have some of the properties of the antibodies then newly discovered by Behring and would be able to find and destroy a target in the body which was responsible for a particular disease.

At a recent meeting* (part of the *Quo Vadis?* series) progress towards the fulfilment of Ehrlich's dream was considered. The aim of the organisers was to assemble a group of membranologists and cancer chemotherapists, immunologists, experimentally orientated clinicians and toxinologists who might collectively constitute an engineering team to consider the numerous component problems of the overall design.

The basic philosophy is to have a carrier molecule, of which the function is selectively to bind to the target, covalently attached to a 'warhead' chosen to kill or modify the target cells. Carrier molecules could be antibodies, lectins, peptide hormones or any other molecule which tissues take up selectively but consideration was limited to the use of antibodies and, in particular, monoclonal antibodies.

There was agreement that the homogeneity afforded by monoclonality was advantageous but restriction of the number of sites to which such a highly specific antibody could bind was seen as a potential problem. Localisation is, however, fundamental to success and it was demonstrated by Mach that localisation of anti-CEA antibodies onto human colonic carcinomas in either mice or men could sometimes be demonstrated (Mach *et al. New Eng.J.Med.* **303**, 5; 1980). In well differentiated tumours it tended not to occur, presumably because of difficulties of access to the luminal (antigenic) surface of the malignant ductal epithelium. It is not known whether the larger molecular weight conjugates would do any better but common sense seems to indicate that they would not. The obstacle is not, however, insuperable in well delineated tumours as various devices could be employed to improve local blood flow or to promote extravasation. Neither is it impossible that IgE anti-tumour antibodies could be used as systemic promotors of extravasation. Alternatively IgA antibodies might be transported across the epithelial cells to the target surface as long as the mechanism for secretory piece recognition could operate.

Warheads so far utilised include small molecular weight standard chemotherapeutic drugs or high molecular weight toxins or their subunits. Enzymes have also been employed.

Toxic agents of small molecular weight have limited effects because relatively few molecules can be bound to the carriers without reducing their cell binding capacity. This problem might be circumvented by binding relatively large numbers of drug molecules to a macromolecule such as poly L-glutamate (Rowland *et al. Nature* **255**, 487; 1975) or dextran (Rowland *Eur.J.Cancer* **13**, 593; 1977; Hurwitz *et al. Int.J.Cancer* **21**, 747; 1978) and then attaching the complex to the carrier with target specificity. It emerged incidentally that merely binding daunomycin to dextran reduced the toxicity of the drug although without apparent loss of therapeutic potential. The mechanism involved here probably involves a slower release of the drug over a long period of time rather than an alteration in its pattern of distribution.

Excitement at the meeting was caused by the description of antibody-toxin conjugates or 'immunotoxins' as Jansen and his colleagues (Centre de Recherche Clin-Midy, Montpellier) call them. The toxins, exemplified by diphtheria toxin, abrin and ricin, consist of a B piece which has a receptor for the cell surface and which may mediate the transport into the cytoplasm of the A piece (Olsnes *et al. Nature* **249**, 627; 1974). The A pieces are inhibitors of protein synthesis which can in various ways interfere with messenger RNA translation in cell free systems and which are presumed to do the same in the cytoplasm of cells to which they gain entry (Pappenheimer *Ann.Rev.Biochem.* **46**, 69; 1977).

The means of linking together the toxin and carrier was discussed. Covalent binding of some sort seemed indicated and various chemical strategies for achieving this were considered. Gros (Clin-Midy) reported that a disulphide bond, such as exists naturally between A and B chains of toxins, was needed to link together ricin A chain and immunoglobulin in a form that could express cytotoxicity in tissue culture. The stability of the SS link *in vivo* was questioned and dissociation of the conjugate may explain the lack of impressive therapeutic results with A chain conjugates in animals. Edwards (Wellcome Foundation, London) reported that use of a disulphide bond to attach intact abrin to anti-lymphocyte globulin yielded a conjugate less capable of inducing immunosuppression in mice than a chlorambucil-linked conjugate.

The membranologists, particularly Blobel (Rockefeller University), were of the opinion that the only molecules that would be able to cross the membrane barrier would be those which had the appropriate translocation signals. The B pieces of toxin molecules have been widely regarded as possessing the signal for translocation of the A chain from outside the cell into the cytosol. This may also be true of diphtheria toxin since conjugates with the isolated A chain have lacked potency (see, for example, Gilliland *et al. Proc.natn.Acad.Sci. U.S.A.* **75**, 5319; 1978; Chang *et al. J.Biol.Chem.* **252**, 1515; 1977). But it seems from the work of Jansen (Jansen *et al. Immunol.Lett.* in the press) and Thorpe (Institue of Cancer Research, London) and others (Uchida *et al. J.Biol.Chem.* **255**, 6687; 1980; Yamaguchi *et al. J.Nat.Cancer Inst.* **62**, 1387; 1979) that conjugates with ricin A-chain or A-piece like materials such as gelonin (Stirpe *et al. J.Biol.Chem.* **255**, 6947; 1980) can also be profoundly and specifically cytotoxic. Antibody molecules, like many if not all other secreted proteins, are thought to lose their leader sequences at the time of their translocation into the vesicle from which they will be shed. Thus it must be considered that certain A pieces, with antibody-conferred capacity to bind to cells, may themselves have the additional information necessary to be translocated

---

*The meeting held in Montpellier, 17th July, 1980, was a satellite symposium of the International Congress of Immunology in Paris.

Tony Davies is Professor of Immunobiology, Institute of Cancer Research, London.

0028-0836/81/010012-02$01.00

into the cytoplasm without a mediating B-piece.

The technology of antibody-toxin conjugates is developing at a time when our understanding of plasma membranes is increasing rapidly. At present we simply do not know exactly how or why toxins gain entry to cells although we can make informed guesses. The consequences of binding of antibody to cell surfaces are also only known in a limited number of circumstances, so it is very difficult to predict what will happen to an antibody molecule coupled to a toxin.

The overall impression was that the new breed of conjugate therapists should increase their efforts to attack targets *in vivo*, in experimental models, perhaps taking into account the pattern of potential clinical usage. The existing results are promising and the answer to the question 'Quo Vadis?' seems to be 'forward'. □

Macroevolution

from John Maynard Smith

In October 1980 a conference was held in Chicago on "macroevolution". At the meeting, some 150 palaeontologists, anatomists, taxonomists, embryologists and geneticists debated the claim that the large-scale features of evolution cannot be explained as a mere extension of the microevolutionary changes within populations which are studied by geneticists.

The main proponents of this view have been Niles Eldredge, Steven Gould and Steven Stanley. The minimum claim is an observational one, that the fossil record does not show a gradual transformation of populations, but instead shows 'stasis' — morphological constancy of species over millions of years — and 'punctuation' — the sudden appearance of new species.

To establish this assertion would require measurements of the distinguishing characteristics in populations from successive strata. An odd feature of the meeting was that no data of this kind were presented. Stanley's[1] book, *Macroevolution*, argues for punctuation mainly on the basis of data on the durations of taxonomic groups (species, genera etc.) in the fossil record. The trouble with this is that one does not know whether one is studying the habits of taxonomists or of evolving taxa. The morphometric data he presents (notably, Gingerich & Simons' data[2] on adapid mammals) show gradual transformation rather than stasis and punctuation. The claim, therefore, appears to rest on the strong impression felt by some, but by no means all, palaeontologists; it may be true, but needs to be shown to be true.

One study known to me which does support the punctuationist view is that of Williamson[3] on 21 species of fresh-water molluscs in the lake Turkana region of Africa. Over the last 5 million years population measurements show only small changes in 16 of the species. The other five change little for most of the time, but rapidly during a period of perhaps 50,000 years when the lake was isolated. The derived species later went extinct and were replaced by the original forms, presumably by immigration from outside. This example ilustrates a semantic difficulty which often confused the conference. A new species arising in 50,000 years (2m of sediment) is sudden to a palaeontologist but gradual to a geneticist. As a geneticist I was left with the impression that the 'sudden' appearance of new species called for no new explanations, but that stasis may do.

Punctuations have made a further deduction from the supposed pattern of stasis and sudden change. This is that long-term trends in evolution are the consequence of 'species selection'. According to the hypothesis of species selection, the species is the 'unit of selection', which either reproduces (that is splits into two species) or dies (that is goes extinct); this is in contrast to the more usual view, according to which the individual is the unit of selection. It is proposed that some lineages show greater speciation rates and/or lower extinction rates than others, and that therefore there is a trend towards those characteristics responsible for rapid speciation and low extinction. Some of the most interesting data presented concerned associations between species' traits (for example, sedentary versus planktonic larvae) and rates of speciation.

The deduction that species selection is the sole cause of trends rests, however, on the further assumption that punctuational changes are as likely to occur in a direction opposite to that of the trend as parallel to it. If this assumption is not true, then the overall trend of evolution would be the sum of the punctuational changes, each one of which would be the result of selection at the individual level within the populations. The punctuational patterns of change would then not imply any new evolutionary mechanism. It would merely show that evolution occurs not at a uniform but at a very varied rate. This would hardly be new; it was in fact the main theme stressed by G.G. Simpson, one of the founders of the 'modern synthesis'.

There is no evidence to show that the direction of punctuational changes is random with respect to major trends, and it seems implausible as a generalisation. It implies, for example, that in the evolution of mammals from reptiles, newly derived species, when compared to their ancestors, were as likely to have more outwardly directed femur and humerus than the opposite, and similarly a shorter secondary palate, a more complete dermal covering to the skull, more bones in the lower jaw, and so on for all the other trends.

Stasis, punctuation and species selection constitute the minimum claim of the new school. Further claims are that the reason for stasis lies in 'developmental constraints', that the origin of new species requires 'hopeful monsters' (large morphological changes arising from a simple genetic change) and/or 'genetic revolutions' (major reconstructions of the genome, perhaps occurring in a small population), and that the fundamental structures, or bauplans, of major taxa represent a set of possibilities constrained by the laws of development; in constrast, Darwin argued that the structures common to a taxon exist because they were present in the ancestors of that taxon, and that they arose in those ancestors as naturally-selected adaptations to particular ways of life.

Thus there are two possible ways of explaining stasis; normalising selection (that is, selection favouring the typical members of a population relative to the extremes) or developmental constraints. Species did not change either because deviants were weeded out, or because the developmental system prevented them from arising. Clearly both effects operate. The crucial test of their relative importance lies in a study of geographical variation in existing organisms. Thus it seems to be the case that species with a wide geographical distribution often display a range of forms such that the most extreme differ from one another as much as do sympatric species. If so, it is hard to see how developmental constraints can account for stasis. There is a clear need to compare patterns of spatial and temporal variation in the same groups.

Is there evidence that macromutations have been important in speciation? The best way to find out seems to be a genetic analysis of hybrids between morphologically distinct species. One example quoted in Chicago concerns hybrids between *Drosophila sylvestris* and *Drosophila heteroneura*[4,5]: the latter has projecting eyes which are among the most striking morphological variants in the genus. Analysis shows that the difference

---

John Maynard Smith is Professor of Biology at the University of Sussex.

1. Stanley, S M. *Macroevolution. Pattern and Process* (W H Freeman, 1977)
2. Gingerich, P.D. & Simons, E L *Univ Michigan Mus. Paleont. Contrib.* 24, 245 (1977)
3. Williamson, P. thesis, Univ Bristol (1980)
4. Val, F C. *Evolution* 31, 611 (1977).
5. Templeton, A R. *Evolution* 31, 630 (1977)

0028-0836/81/010013-02$01 00

in eye shape is caused by genes at at least eight loci. It seems that no macromutation was involved in this case. I know of no example pointing the other way, but shall not be surprised if one crops up.

Genetic analysis cannot, of course, be applied to the more extreme differences between bauplans, so it is still possible to argue that the origins of major taxa require special genetic events. However, no-one at the conference was prepared to challenge Stebbins's remark that the most striking morphological novelty among the higher plants was the flower head of *Zea mays*, which is largely the result of selective breeding under domestication and is known to be under the control of many genes with small effects.

It will be apparent that this report has been written by a geneticist, not a palaeontologist. Despite my reservations, I found the meeting immensely stimulating. It can only be good for evolutionary biology that people from such different disciplines should meet, talk and, occasionally, listen. It was only to be expected that there would be much misunderstanding, confusion and even indignation. I hope we shall meet again in a year or two, and that next time the confusion may be slightly diminished. □

Why is a plant where it is?

from Peter D. Moore

ONE of the simplest questions to ask, yet one of the most difficult to answer is why a particular individual of a given plant species should be found growing at a particular location. The answer may be sought in its physiological requirements and the degree to which they are satisfied by the physical framework of that spot; or in the complex of biological interactions with other organisms, competitors, predators, parasites, and so on; or in the tangled web of historical events which has moulded both of these former factors; or, finally, in the sheer chance which brought a propagule to that site and which left it untouched by random catastrophe. The interaction of all these forces permits no easy answer to the apparently simple question.

Any attempt must seek first to locate the physical environmental stress and physiological sensitivity which, in combination, allow the identification of a limiting factor. This factor can be regarded as a necessary resource in such short supply that any increase in its availability will result in an increase in the performance of the individual or density of the population. But the model, stated in these terms, needs to be flexible enough for such factors as temperature at specific times of day or year to be regarded as a resource.

One may examine any aspect of the plant's physiology, or any time in the plant's life history, to find where resource restriction meets the plant at its most vulnerable point and prevents further reproduction, growth, or even survival. Perhaps the most obvious place to start is at the beginning and to define those resource limitations which restrict the germination of the seed and thus determine whether the plant is to be given the opportunity to meet further tests. Thompson (*Ann. Bot.* **34**, 427; 1970) has approached from this direction in his studies of various species of Caryophyllaceae. He devised a metal bar along which a temperature gradient was maintained and observed the germination characteristics of each species in relation to temperature, and found that their requirements correlated well with the biogeographical characteristics of each species. For example, a species (*Silene secundiflora*) from the Mediterranean germinated at low temperature, this being advantageous to species in which autumn germination means a mild, damp period before meeting the unfavourable drought of the following summer. A more northerly species, on the other hand (*Lychnis flos-cuculi*), had a higher temperature germination requirement; it thus remains dormant through the winter after seed set and germinates in the spring.

Subsequently Thompson (*Ann. Bot.* **39**, 1; 1975) has examined in detail the geographical variation between populations of a single species, namely *Silene dioica*. He has found considerable variation both within and between populations and also a different response among after ripened seeds from those which were fresh. The populations from regions with milder winter contained more individuals which were capable of germination at low temperature, thus confirming the trend indicated by the previous experiments.

Silene dioica, the red campion, is a woodland herb, widespread in Europe, which, in the western, deciduous woods of Britain and northern France is often accompanied by the bluebell, *Hycaninthoides non-scriptus*. The seeds of this species were also subjected to the attentions of Thomson and Cox (*Ann. Bot.* **42**, 51; 1978), and on the basis of their experiments they propose that the seeds germinate naturally at about 11°C, in the late autumn and that they therefore overwinter as seedlings. This obviously leaves the species somewhat vulnerable to winter frost; its distribution appears to correlate well with the isotherm of the main daily maximum temperature of the coldest winter month at 0°C, and also with a mean daily maximum of the warmest summer month at 25°C. Its sensitivity to winter frost and summer drought account for its western distribution and its confinement to the woodland habitat, where it is protected from both. The survival of bluebells in open habitats in the extreme west of Britain (often as a relict of former woodland) may be a reflection of the mild, oceanic conditions prevailing there.

The examination of seed from one further woodland species, the wood millet grass, *Millium effusum*, has shown a rather more complex pattern of reaction. Thompson (*Ann. Bot.* **46**, 593; 1980) found that seed germinated immediately after harvesting had an optimum temperature requirement of 16°C but, after storage for 2.5 months, the optimum temperature for germination was 21°C and germination took place over a wider temperature range. The breadth of tolerance increased yet further after 10 months. Also, germination was faster after a period of storage but, rather surprisingly, a higher temperature (26°C) of storage reduced the subsequent germination rate. The ecological significance of this high temperature inhibition of the after ripening process is difficult to discern, but Thompson suggests that it could reduce the chances of an ill-advised, rapid response to mild conditions soon after shedding.

Millium effusum has a wide distribution in Europe and, in view of the late flowering and fruiting in the north of its range (late August in areas to the north of the Baltic), temperatures will already be below those needed for germination, which is thus delayed until spring. In the south (southern France and Iberia), seed is shed in early summer, so it experiences a hot, dry period followed by a warm, moist one. This results in the bulk of the seed germinating before winter arrives. In between these two extremes, the complex microclimatic variations caused by site factors (microtopography, aspect, etc.) and fluctuations in weather conditions, can result in a range of responses. Indeed, the plasticity of this species may be a key to its success.

Thompson's germination studies illustrate the possibilities which exist for pinpointing the limits of a plant's capacity to meet the stresses of the environment, but seed germination is only one of many critical phases in the plant's life cycle. Also, the fact that one factor may limit a species on a continental scale, does not preclude the possibility that other factors will limit its local distribution pattern. At the level of the individual one must still account for the part played by chance. At this scale is vegetation simply, in the words of Gleason, a "fortuitous juxtaposition of plant individuals"?

Peter D. Moore is in the Department of Plant Sciences, University of London King's College.

0028-0836/81/010014-01$01.00

Galileo would have been pleased

from J. E. Guest

THE discovery that Io, the Moon-sized satellite of Jupiter, has active volcanoes throwing plumes as much as 270km above the satellite's surface was one of the exciting revelations made by the Voyagers I and II spacecraft as they sped past Jupiter and its attendant satellites in March and July 1979. The four largest satellites in the Jupiter system, named Io, Europa, Ganymede and Callisto, were discovered by Galileo in 1610. For three centuries after their discovery little more was found out about them but, during this century, it became abundantly clear, even from Earth-based astronomical observations, that each satellite has distinctive characteristics.

Although classified as satellites, the Galilean moons are planetary bodies in their own right and of similar size to the Moon and Mercury. Io, the innermost of the satellites, was thought to be the most peculiar. Its density (3,500kg m^{-3}) is similar to that of terrestrial rocks and its surface is characterised by high reflectivity and a striking reddish colour. Furthermore, Io is constantly emitting atoms from its surface, leaving a trail of ionised sulphur, potassium, hydrogen and sodium behind it to form a torus-shaped atmosphere in its orbital path. Telescopic observations of the next satellite outwards, Europa, showed that it too has a density (3,100kg m^{-3}) consistent with it being made primarily of rocks. Its density is slightly less than that of Io implying a higher water content. Whereas Io is somewhat larger than the Moon, Europa is almost the same size.

The way light is reflected from Europa and the next satellite out, Ganymede, suggests a covering of water ice on their surfaces. It is thought that Europa is between 50 and 100 percent frost and ice, and Ganymede between 20 and 60 percent covered by fresh ice. Ganymede has a much lower density than Io or Europa, (1,900kg m^{-3}) implying a high content of water and ice. Callisto, on the other hand, which is of comparable density (1,800kg m^{-3} from new Voyager data) to Ganymede, has a low albedo and any ice at the surface must be 'dirty' containing fragmental rocky material.

Thus from the limited pre-Voyager astronomical data it was clear that the internal structures and compositions, and the processes that had moulded the surfaces could be very different from those with which we have become familiar on the terrestrial planets. From previous studies of the Moon, Mars and Mercury, a familiar pattern of surface history for solid planetary bodies has emerged. The early surface history was dominated by impact craters, ranging from the minute to large basins of 13,000k diameter or more. The surface expressions of any other processes operating at the same time, such as volcanism, were destroyed. For the Moon at least, this period of bombardment ended around 4×10^9 years ago. There then followed for each of these planetary bodies a period when volcanism dominated to give on the Moon, for example, the extensive dark marial plains, consisting of vast sheets of lava. On Mars, the volcanic process was more varied giving large volcanic constructs such as the huge mountain of Olympus Mons. Tectonic processes produced patterns of faults, the most complex structures being on Mars. As well as these internal processes modifying the surface Mars was further modified by the effects of water and wind erosion as well as the effects of sub-surface ice. But even these processes occurred relatively early, the major events having been completed $2\text{–}3 \times 10^9$ years ago.

The pictures returned by the Voyager spacecraft show surfaces much stranger than had been predicted.

Whereas for bodies in the inner Solar System virtually the whole record of surface history has been preserved, the surfaces of the Galilean Satellites have each preserved different periods in their history and each is dominated by specific processes.

The outermost satellite, Callisto, has a surface dominated by impact craters and basins, and presumably has the oldest preserved surface of the Galilean satellites. The high content of ice and water suggests that there was little modification of the surface by internal activity following the period of impact bombardment. Ganymede on the other hand has a higher rocky content and as well as impact craters a complex pattern of tectonic features which have strongly modified the surface since the period of heavy bombardment can be seen. The evidence suggests that during the period of bombardment the mechanical crust of Ganymede was probably only a few tens of kilometres thick and there then followed a period of rapid thickening and stiffening of the crust resulting in the tectonic patterns now observed.

When we examine the surfaces of Europa and Io, we can see they have quite different histories. On Europa only a few impact craters have been identified suggesting that the present surface was formed after the period of intense bombardment; on Io, volcanic resurfacing has obliterated the craters formed earlier. Thus, although on Ganymede internal activity has been important, its effects on the surface were concluded early in its history, whereas for Europa and Io, internal mechanisms have continued, and, in the case of Io are still operating.

Models for the internal constitution and thermal history of the two inner Galilean satellites have concentrated on two energy sources, one deriving from tidal interaction between Jupiter and the satellites and the other involving decay of long lived isotopes within the rocky interiors. Models involving these two processes depend very much on the satellites' origin, the timing of events and the composition and degree of differentiation.

In their remarkably well timed paper, Peel *et al.* (*Science* **203**, 892; 1979) produced a model in which there was intense heating in the interior of Io by tidal dissipation, anticipating the Voyager results of active volcanism on Io just a few days before it was discovered. Likewise, tidal heating can have played a part in the thermal history of Europa and tidal deformation may explain some of the surface characteristics. The surface of Europa is generally smooth and bright with a strange pattern of intersecting linear features composed of shallow grooves, low ridges and albedo markings. The topographic relief on these features is small and only seen close to the terminator. These features are probably formed by stress in the surface materials, but their orientation has not so far been shown to be uniquely explicable in terms of a particular stress field. Previous models for the internal constitution of Europa suggested that early differentiation resulted in a silicate core surrounded by a liquid water mantle, crusted over by some tens of kilometres of water-ice. However, more recent calculations suggest that solid state convection in the crust probably caused progressive freezing of the liquid mantle and it now seems unlikely that tidal heating could have maintained the liquid state.

The Voyager discoveries have prompted a reconsideration of this type of model. It is possible that the six percent by weight water component (derived from density considerations) is not all taken up as free water or ice at the surface, is instead contained in hydrous silicates in the rocky part of the planet, leaving only a thin ice outer shell of a few kilometres thickness (Ransford, *et al.* this issue of *Nature* page 21). However, such a model imposes constraints on the time of formation of the surface ice. It appears that Europa's surface consists of relatively 'clean' ice, although in some areas a brownish hue to the surface may suggest contamination by rocky material. Assuming that Europa, like the outer Galilean satellites, was bombarded in its early history it is

---

J.E. Guest is a lecturer in Planetary Science in the Physics and Astronomy Department, University College London.

0028-0836/81/010015-02$01.00

inevitable that, with a thin icy crust underlain by rock, impact excavation would have brought up rocky material, mixing it with the surface ice: the present ice layer must therefore have formed after the period of heavy bombardment. With a thicker ice (or water and ice) outer shell, impact need not have penetrated to the rocky layers and the effects of impact might have been removed by cold viscous flow in the ice crust. Alternatively, early formed craters might have been eroded by the sputtering effect of charged particles capable of removing tens of metres, or even kilometres, from the surface (Smith *et al. Science* **206**, 927; 1979). A thin icy crust must either have formed after the end of bombardment, or have remained in a liquid state, allowing denser silicate material to settle, until after the bombardment period.

The origin of the linear pattern on Europa is particularly enigmatic. Various proposals have been put forward including crustal expansion owing to freezing of an early ocean, global expansion as a result of dehydration of the interior, and tidal deformation. The complexity of the pattern made up of features of different morphology is difficult to interpret in terms of mechanism and indeed more than one process may have contributed to the surface features. At present, interpretation of the surface of Europa is model-dependent; a situation which is perhaps not surprising on a planet whose geology is unlike anything in our previous experience. □

Finding nitrogen fixation genes

from John E. Beringer

THE enzymic conversion of atmospheric nitrogen to ammonia is carried out only by prokaryotic microorganisms. The structure and properties of the enzyme complex used, nitrogenase, are very similar in all species that have been studied.

Nitrogenase consists of two main components. One is an iron protein, composed of two identical subunits; it binds MgATP and transfers electrons to the other component which reduces the substrate. The second component is a molybdenum-iron enzyme containing four proteins which can be separated into two distinct sub-units, giving the enzyme an $\alpha_2\beta_2$ type of structure. All these proteins are irreversibly denatured by oxygen. A further striking similarity between nitrogenases is that a functional enzyme can be constructed by mixing iron and iron-molybdenum proteins from different genera; for example *Klebsiella* and *Azotobacter*, though some combinations such as *Klebsiella* and *Clostridium* are inactive.

Over the last few years there have been dramatic advances in our understanding of the genetics of nitrogen fixation (Nif). This has been due largely to the fortunate finding that *Klebsiella pneumoniae nif* genes are linked to each other and could be manipulated in *Escherichia coli*. As a result of extensive mutant and complementation studies we know that there are 17 *nif* genes in seven or eight operons. Three genes code for nitrogenase proteins; *nif* H for the iron protein and K and D for the different subunits of the molybdenum-iron enzyme. Functions have yet to be assigned to some of the other genes.

Parallel genetic studies of *nif* genes in other nitrogen fixing species have been restricted by limitations in our ability to do the necessary genetic manipulations or, in the case of the genetically amenable *Rhizobium* species, our inability to obtain nitrogen fixation in pure culture. However, using *Rhizobium meliloti* as an example, Ruvkun and Ausubel in this issue of *Nature* (see page 85) have shown that one can take advantage of the conservation in DNA sequence of *nif* genes to promote site-directed mutagenesis and localized mapping in prokaryotes.

In 1979 Nuti *et al.* (*Nature* **282**, 533) reported that DNA encoding the *K. pneumoniae* iron and molybdenum-iron proteins hybridized with plasmid DNA from *Rhizobium*, though DNA encoding most of the remaining *Klebsiella nif* genes did not hybridize. In January this year Ruvkun and Ausubel (*Proc. natn. Acad. Sci. U.S.A.* **77**, 191) and Mazur, Rice and Haselkorn (*Proc. natn. Acad. Sci. U.S.A.* **77**, 186) reported further evidence for the conservation of *nif* genes. Ruvkun and Ausubel looked at 19 strains of different nitrogen fixing prokaryotes, including Gram-positive and Gram-negative bacteria, cyanobacteria and the actinomycete *Frankia*. All showed DNA hybridization with the fragment of *K. pneumoniae nif* DNA carrying genes K D and H used by Nuti *et al.*, whereas no hybridization was observed with controls representing ten different non-nitrogen fixing species or with DNA carrying the other *nif* genes. Mazur, Rice and Haselkorn confirmed that hybridization could occur with three strains of cyanobacteria representing two different species. They further showed by heteroduplex studies of hybridized DNA that the genes were not contiguous cyanobacteria strains. It is unclear from the published data for *nif* gene hybridization whether or not all three structural *nif* genes (H D and K) hybridize in all these species. Hopefully this will be resolved in the near future.

Another problem which should be resolved soon is the determination of the extent of DNA sequence homology for these genes in different species. The DNA sequences for the *nif* H proteins of *Anabaena*, (Haselkorn pers. comm.) and *Klebsiella* (Ausubel pers. comm.) are now available. It will be interesting to see how highly conserved the sequences of the different genes are and whether or not the data support the idea that the enzyme has been conserved over a long evolutionary period, or that nitrogen fixation is a relatively new function which has been transferred between different bacterial genera.

These DNA hybridization studies have been of very great value for comparative studies of *nif* genes in different families. As described above, they can be used to determine whether *nif* genes are chromosomal or plasmid borne, whether they are linked to each other and as a method for isolating *nif* DNA for studies of its expression in different hosts or for further genetic studies. It is these last two uses of the *nif* hybridization work which are discussed by Ruvken and Ausubel in this issue of *Nature*. They show, as one would expect, that the *K. pneumonia nif* DNA can be used to isolate the corresponding *Rhizobium* genes and introduce them into *Escherichia coli*. Of much greater importance scientifically is the demonstration that this procedure can be used to induce mutations in the *Rhizobium nif* DNA within *E. coli* and then return the defective gene to its original host where recombination can be forced so that the defective gene is introduced into the chromosome replacing the existing functional gene. In this way mutants of *Rhizobium*, or potentially any other nitrogen fixing prokaryote, which are defective in known *nif* genes can be induced.

Once such a mutation has been induced it is possible to re-isolate DNA fragments extending from this region and induce mutations in them. By knowing the extent of the overlap with the original fragment a series of mutations can be induced at known distances from the original mutation by repeated rounds of mutagenesis and replacement of the corresponding host DNA. While the value of this technique for studies of genes linked to a particular *nif* gene is obvious it is also equally valuable for all species where cloned DNA fragments are available and there is a method for re-introducing the mutated DNA into the host. Thus we now have a technique which could be used to prepare maps of chromosomes or plasmids in a wide range of species of microorganisms.

J.E. Beringer is in the Soil Microbiology Department, Rothamsted Experimental Station, Harpenden, UK.

0028-0836/81/010016-01$01.00

ARTICLES

Lineament and polygon patterns on Europa

David C. Pieri

Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103

A classification scheme for the lineaments and associated polygonal patterns seen on the surface of Europa is presented, and the polygon frequency distribution discussed.

THE surface of Europa is dominated by lineaments intersecting to form polygons and by complex systems of low ridges which often exhibit cycloid planimetric patterns[1,2]. Polygonal or nearly polygonal patterns are common in nature and are often associated with stress-relief fracturing. The interpretation that the albedo lineament patterns are associated with fracturing at different scales is used. Two major pattern groups are of interest: (1) roughly equant polygons formed by the surface expression of fractures formed in isotropic stress regions, both tensional and compressional (for example, shrinkage cracks in mud, cooling cracks in basalt, patterned ground, fracture patterns over salt domes); and (2) complex trapezoidal patterns formed by the intersection of faults or joints under directed stress.

The existence of either of these pattern types has implications for the state of stress under which each type formed. This article presents a classification scheme for the lineaments and associated polygonal patterns seen on Europa, and reports on the polygon frequency distributions. Possible fracture mechanisms and scenarios for the formation of lineament systems on Europa are examined in detail elsewhere[3–5].

Lineaments

Figure 1 is a low resolution sketch map of the global lineament patterns on Europa. The general orientation of lineaments along great circles is apparent and has been mentioned previously (ref. 1 and Plescia personal communication). Lineaments are seen less frequently in low albedo regions, although this may be due to loss of contrast across dark areas.

Figure 2 is a high resolution (4 kilometre per line pair (km/lp)) mercator projection mosaic. Figure 2 shows that at small scale the lineaments form extremely complex patterns. Some lineaments exhibit relief while others seem to be primarily albedo features. Most dark lineaments possess a white medial stripe and often, when dark lineaments lie across low albedo terrain, only the white interior stripe is visible across the low albedo patch. The first step to try to understand these enigmatic features was devising a preliminary classification system which is outlined in Table 1 and Fig. 3. Although this classification is rudimentary it allows different styles of lineaments to be discussed. The lineaments have been divided on the basis of albedo, morphology, orientation, and characteristic geometric style into eight groups. However, this analysis is based entirely on the limited set of Voyager 2 images taken from a distance of ~100,000 km and the best resolution of the images is ~4 km. Stereo coverage, while available, is marginal and of low resolution. Coverage at 4 km resolution is available only for ~25% of the planet, the remainder of the surface is covered at resolution at least an order of magnitude worse. Although this limits the analysis the classification scheme is, nevertheless, representative of the region for which coverage is available.

Figure 3 shows the various lineament geometries and Table 1 gives type descriptions and locations. These lineaments on Europa are clearly of several length scales and there are at least two basic types. Flexi (Type 6) and scarps (Type 8) seem to be topographic features with little or no albedo contrast to the surrounding terrain. The flexi are global features, and we do not know the extent of the low scarps. The linea appear to be primarily albedo features. Type 3 with white centre striations appear to have some positive relief although Soderblom (personal communication) reports that based on photometric measurements relief may be small to non-existent. Negative relief for one northern linea has been suggested[2]. Lineament Type 3 is clearly global. These are generally large great circles. Type 1, 4, and 7 seem to be regional in extent, whereas Types 2 and 5 may be the result of highly localized processes.

Polygonal areas

All the lineaments mentioned above define a system of polygons varying in size from small (limit of resolution 10 km^2) reticulate polygonal patterns to very large (10^6 km^2) individuals. The statistics of polygonal patterns in nature are well-known[6,7].

For macroscopic geological settings two pattern types are common in terrestrial experience: (1) equant five or six sided polygons; and (2) predominately four-sided polygons formed by faults or joints in complex settings.

The two types have distinct and characteristic frequency distributions for polygons of a given number of sides: the so-called Voronoi distribution (mode 6) is characteristic of polygonal patterns formed under uniform tension or compression; the Poisson distribution (mode 4) being characteristic of polygons formed by random intersections of lines on a plane or by the intersections of random great circles on a sphere[6,8,9] (Fig. 4*a*).

Literature data on several types of terrestrial polygonal patterns were compiled for polygon side-frequency distributions and these are presented in Fig. 4*b*. Slowly cooling basalt is reported to possess equant polygonal patterns[10–12]. Although the most famous columnar basalt occurrences exhibit a roughly Voronoi polygon distribution, they tend to have an excess frequency at the modal value[13]. Further, basalt, cooling in less than ideal conditions for the formation of polygons, tends to have a mode of five rather than the classic six. Rapidly cooling, basalt-glass crusts on lava lakes tend to have a different character[14]. Although they are grossly polygonal and have a large number of five- and six-sided polygons, they have a large number of three- and four-sided members as well. Concoidal fractures are common, and this characteristic undoubtedly plays a part in the observed side-frequency distribution. It is often difficult to define polygonal forms at all under these circumstances. Mud crack networks[6,12,15–19], and tensional fractures over salt domes[20] typically follow a Voronoi distribution as do contractional thrust ridges in playa salt crusts[21]. Ice wedges and patterned ground are also Voronoi distributed[7,22]. The common feature of these examples is their formation in isotropic or near isotropic stress conditions, which is not surprising as the ideal hexagonal pattern maximizes the reduction of strain energy per unit surface area by cracking and is therefore the most efficient pattern given an isotropic stress field (see ref. 7).

0028–0836/81/020017—05$01.00

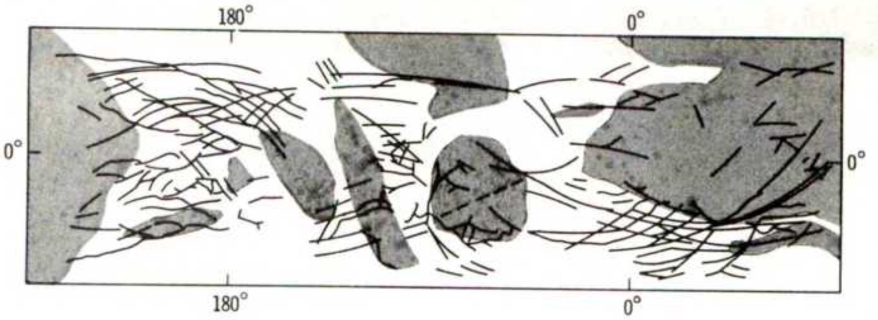


Fig. 1 Global map of lineament patterns on Europa between roughly ±30° latitude. Sub-jovian point on the planet is located at lat 0°, long 0° and the anti-jovian point at lat 0°, 180° long. This map is a cylindrical projection, was compiled from low resolution full-disk images from Voyager cameras. Shaded areas correspond to low albedo regions as seen in Voyager blue-filter images. The absence of many apparent lineaments in these areas may be due to either a lack of contrast or to a lack of lineament development in these areas.

The Poisson distribution characterizes randomly intersecting lines in a plane or on a sphere[6]. Statistics for various fracture and joint patterns are presented in Fig. 4*c* (refs 23, 24). All follow a left-skewed Poisson-type distribution. The most random fracture patterns, as expected, follow the Poisson distribution quite well; however, fracture patterns from strongly oriented stress fields tend towards having orthogonal intersections, thus resulting in left-skewed distribution but with great excess in frequency over the Poisson expectation value at the mode of 4.

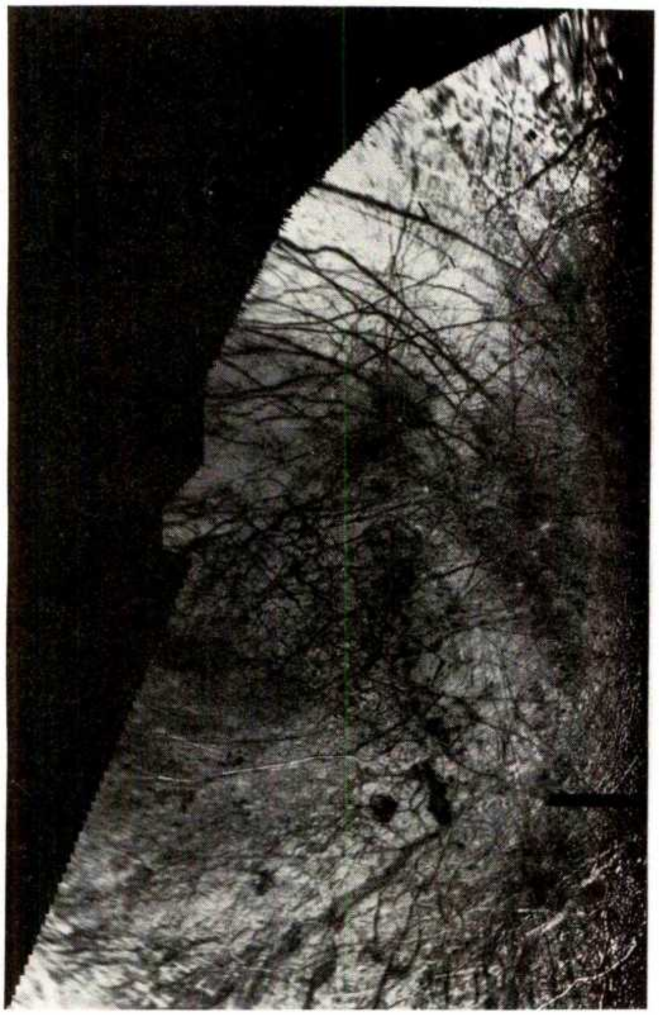


Fig. 2 A computer-enhanced and mercator projected mosaic of Voyager 2 images (resolution up to 4.5 km/lp) taken through the blue filter, in which about 25% of the European surface is shown in this mosaic. The anti-jovian point is near the dark, intensely lineated region at mid-left. Map coordinates of the upper right corner, lower right corner, and lower left corner are respectively: lat +65°, long 130°; lat −70°, long 130°; and lat −70°, long 240°.

If Europan lineaments were expressions of fracturing in the crust of Europa, then it would be logical to compare the statistics of polygons formed by lineament intersections with the statistics for various terrestrial fracture patterns as outlined above. The fundamental distinction between fracture patterns with a left-skewed (Poisson-type) side–frequency distribution and those with a mildly right-skewed (Voronoi-type) pattern is the presence of 'triple junction' fractures. In an isotropic stress environment (tensional or compressional) failure tends to create nonorthogonal intersecting fractures with intersection angles of about 120° (see ref. 21 for examples of compressional isotropic failure). These triple junctions tend to occur as part of five- and six-sided polygons. The more isotropic and uniform the stress, the more likely it will be that six-sided polygons form. For Europa it is important to determine whether structure exists in the pattern of myriad intersecting lineaments. Preferred orientations or coherent structure in the lineament patterns has implications for the organization of stresses in Europa's crust, if the fracture hypothesis is correct.

An apparently qualitative coherence to at least part of the lineament pattern has been reported[1,2,25,26]. Parmentier *et al.*[27] attempted to analyse the orientations of various lineament groupings and compare these trends with stress fields predicted by tidal stress models[28]. We have analysed the polygonal patterns[29] to find out whether there are statistical differences in the side–frequency distributions between polygons resulting from the intersections of the various lineament types, and if so, what those differences imply about the state of formational stress in the surface of Europa.

Preliminary analysis of three major polygon patterns are reported here:

(1) The polygons formed by the intersections of Type 3 lineaments.

(2) The polygons formed by the intersections of Type 1, Type 2, Type 4 and Type 5.

(3) The roughly equant polygons formed by the intersections of Type 1 and Type 2 lineaments.

The side–frequency distribution for polygons formed by Type 3 lineaments (Fig. 4*d*) are similar to those found in complex faulted terrain on the Earth, particularly in areas where stresses are well oriented. They have a mode of 4, are left-skewed and exhibit a Poisson-like frequency distribution. The maximum principal stress is apparently oriented roughly north–south as can be seen from lineament orientation[27]. Nothing can be inferred from the plan pattern alone about the sign of the stress in this case. Some models, however, suggest that it may be compressive[27,28].

Polygons formed by the intersections of Types 1, 2, 4, and 5 lineaments occur around the anti-jovian point on Europa. Their overall distribution is not grossly different than the polygons formed by Type 3 lineaments with one important exception: the number of five- and six-sided intersection polygons mapped for the Type 1, 2, 4 and 5 lineaments exceeds those seen in the Type 3 distribution by a factor of 2 and a factor of 12, respectively. That is, there is a significant shift in the distribution to the right (Fig. 4*d*).

The 'wedge-shaped lineaments' referred to elsewhere[2,25,26], and called Type 2 lineaments here, have been cited as evidence for a nascent type of plate tectonics on Europa[26]. The presence of numerous 'triple junctions' or three-cornered fractures associated with the intersections of Type 1 and Type 2 lineaments supports this line of speculation. Alternatively, they may be the result of local expansion/contraction phenomena. Regardless of their potential connection with possible plate movements, the Type 2 lineaments, when they are mapped separately from other lineaments, have a grossly equant polygonal pattern (Fig. 5). The side-frequency distribution associated with the polygons formed by Type 1 and 2 (Fig. 5) lineament intersections is shown in Fig. 4d. It appears to be much more right-skewed with a mode of 5, similar to the Voronoi-like distributions seen in tensional

Table 1 Descriptions and locations of lineament types

Lineament type	Location of best examples	Description
1	lat −55° to equator, long 160°	175–500 m long, regular spacing, trend ENE, soft appearance, disrupt and are disrupted by Type 2. Together, Types 1 and 2 form the generally equant equatorial polygonal features discussed in text.
2	lat −50° to equator, long between 160° and 210°	Short wedge-shaped markings; typically straight; form triple junctions; together with Type 1 lineaments they form equant polygonal features around anti-jovian point; darkest of the lineaments in Voyager blue filter; typically trend NW–SE.
3	lat +30°, long 180°	Smooth arcuate features; appear to lie on great circles; longest and most continuous of the linea; white medial stripe; remarkable continuity over large distances; intersect at shallow (≤30°) angles or are parallel; several tens of km wide; cross dark terrains of high and low albedo.
4	two main groups: lat +30°, long 165°; lat −30°, long 160°	Two main groups, but may be ubiquitous in light terrain; faint; reticulate patterns; linear; intersect at high angles; occur in remarkably parallel doublets which maintain constant separation (50 km) over distances of hundreds of km; narrowest of lineaments (5–10 km); subdivide polygons formed by the intersections of Types 1, 2, and 3; seem to cut Type 3 lineaments and are cut by Types 1 and 3; intersect Type 3 lineaments at high angles (80°).
5	lat 0° to −40°, long 165° to 210°	Very short (<10–20 km in length); subdivide equant polygons formed by Types 1 and 2 to the limit of resolution; ubiquitous around anti-jovian point.
6	lat −60°, long 150–210°	IAU *flexi*; complex series of cuspate ridges lying on a small circle crudely centred on the anti-jovian point and the complex of Types 1, 2, and 5 lineaments; appear as distinct ridges with a regular cycloid pattern[2]; one set convex towards the anti-jovian point, the other convex away; arranged in radial pattern around a possible impact crater (may be artefact of position with regard to sub-solar point).
7	lat −40°, long 180–220°	Rare; only two visible examples; may be a modification of Types 1 or 3; extremely high albedo; smooth, continuous and linear; very fresh appearance; not cut by other lineaments; one instance of darkening at intersection with another lineament.
8	lat −65°, long 120°; lat 55, long 120°	Possibly low sun facing scarps of indeterminate heights[2] although Soderblom (personal communication) claims they are albedo markings only; roughly parallel to trend of nearby flexi.

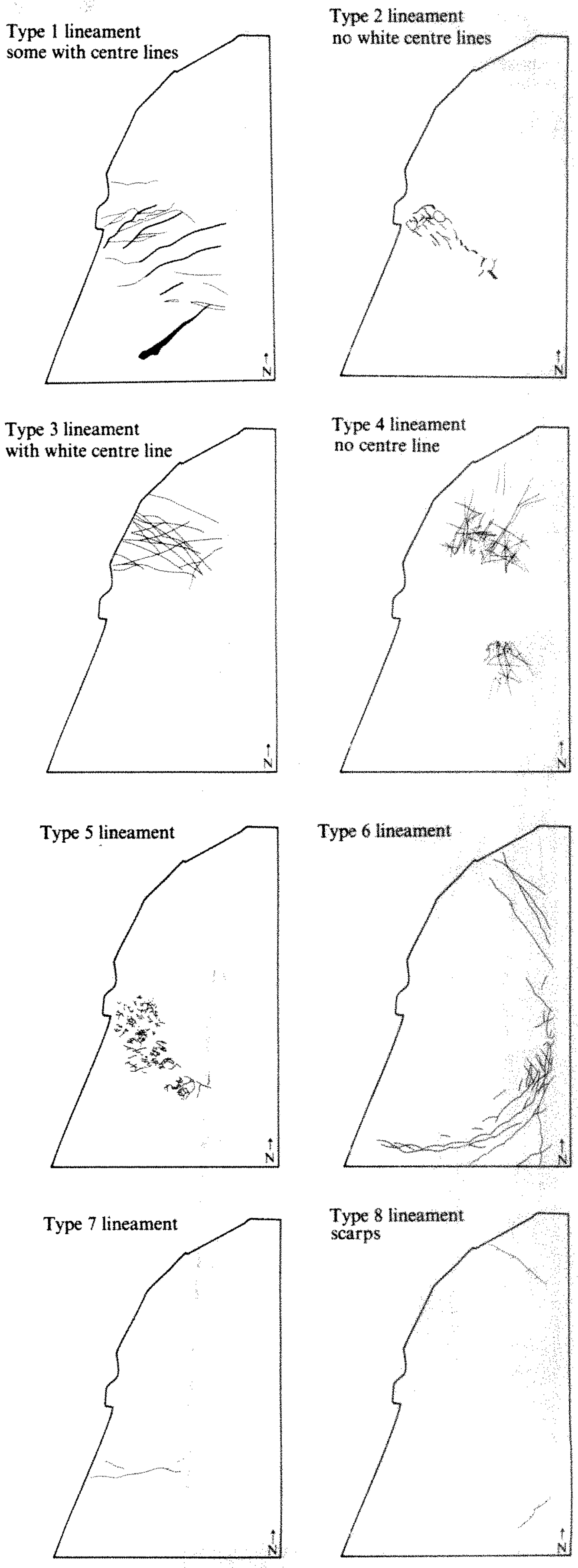


Fig. 3 Sketch maps showing the geographic positions and planimetric morphology of the various types of lineament features observed on Europa. The map outline is the limit of visibility shown in Fig. 2.

fracture patterns. Again, however, the pattern cannot provide information on the sign of the imposed stress; however, isotropic or near-isotropic tension is the predominate factor in terrestrial examples of this style of failure.

Conclusions and discussion

Qualitative and quantitative observations of lineament and polygon patterns on Europa enable several conclusions to be drawn.

First, definite classes of lineaments and their associated polygons have been found on Europa which are separable both by qualitative and quantitative analysis. Second, these classes form distinct groups-often localized, but also some of global character. Finally, these groups have separable geometrical and transectional characteristics, probably related to the stress fields in which they were formed.

These conclusions lead to several speculations. First, the several morphological classes and scales of lineaments may imply different processes acting at different scales, and may also be related to a temporal evolution of the lineament forming processes as suggested by Soderblom[25]. Further, geographically diverse areas on Europa may respond differently to lineament formation, which may be construed as evidence for heterogenity in the European crust. Classification schemes proposed so far[1,2,25,29] indicate a diversity of lineament type and thus, a complex geologic history.

Second, the observations of roughly equant five- and six-sided polygonal patterns near the anti-jovian point argue for a grossly uniform (probably tensional) stress field around the anti-jovian point on Europa. The uniform tension at the anti- and sub-jovian points may be consistent with tidal eccentricity stresses[5], however, preliminary calculations (B. Bannerdt, personal communication) point to very low (<10 bar) stresses from this mechanism. Other models invoking internally generated stresses (such as convection, and dehydration of interior) are being pursued[3,4] and suggest much higher stress levels in Europa's crust. Further, the areal density of the lineament patterns and associated polygons varies with position. Overall, the spacing of tensional polygons varies with material strength and/or rate of stress imposition[7]. Closely spaced reticulate patterns occur at the anti-jovian point and may be indicative of higher imposed stress, of a weaker (or possibly thinner) surface layer, or both.

A more detailed morphological/time-history lineament classification scheme, detailed morphological and geological

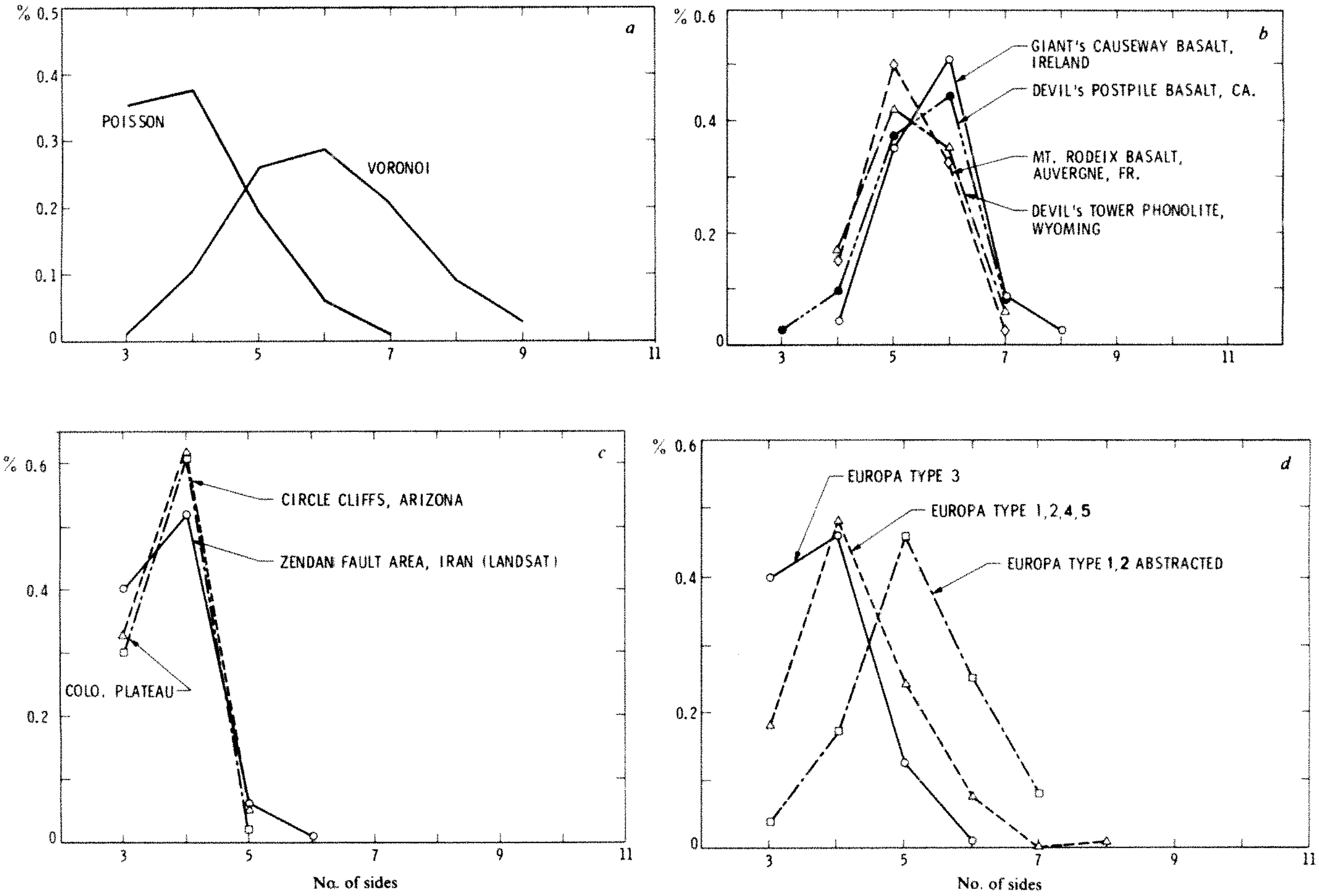


Fig. 4 *a*, Theoretical frequency distributions for polygons of a given number of sides formed by randomly intersecting lines (Poisson) and tesselation polygons formed, for instance by cracking mud (Voronoi) (adapted from Tables 2.1 and 2.2 of ref. 6). *b*, Measured side-frequency distributions for polygons formed by tensional cracking in columnar basalts at well-known localities, compiled from Beard[13]. Note that modes of both 5 and 6 sides are prevelant. The most pristine examples, having formed in presumably ideal conditions (slow cooling rates, few edge effects) have frequency modes of 6. *c*, Side-frequency distributions for intersecting faults in complexly faulted terrain in three areas on the Earth. All show an excess over the random Poisson distribution at the mode. The examples from Circle Cliffs, Arizona and from near the Grand Canyon are in areas where predominate fracture trends are nearly orthogonal, hence the extreme enhancement over the Poisson expectation mode. The Zendan fault area in Iran is part of the Zagros crush zone where faults are predominately parallel to the axis of the collision zone and tend to intersect at shallower angles than the other two examples. Nevertheless, there is enhancement at the mode over the Poisson expectation value, however, it is about 10% less than in the orthogonal systems.) *d*, Side-frequency distributions for polygons formed by various lineament types on Europa.

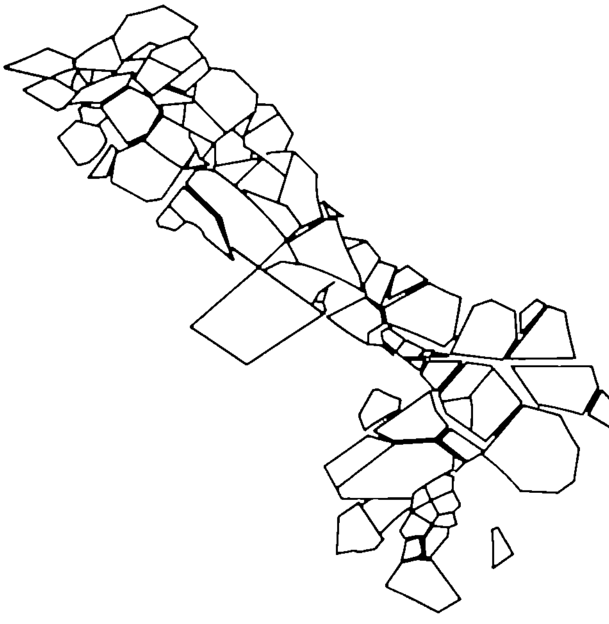


Fig. 5 Sketch map of polygons formed by the intersection of Type 1 and Type 2 polygons on Europa. This region is centred at about lat −20°, long 190°.

studies bearing on lineament formation mechanisms[3,26], and more comprehensive statistical analysis of patterns are needed both to confirm observed pattern structure, and to be used as an alternative and quantitative classification tool. Ultimately the existence of lineament patterns must be accommodated within, and may provide insights for, models of the structure and evolution of both the interior and surface of Europa.

I thank Ken Collerson, Tony Finnerty, Mike Malin, Dennis Matson, Mark Parmentier, Larry Soderblom, and Gary Ransford for useful discussions. This research was carried out while I was a NASA-National Research Council Resident Research Associate at the Jet Propulsion Laboratories of the California Institute of Technology.

Received 5 August, accepted 8 December 1980

1 Smith, B A *et al Science* **206**, 927–950 (1979)
2 Malin, M C *IAU Colloq* No 57 (1980)
3. Finnerty, A A, Ransford, G. A., Pieri, D C. & Collerson, K. D. *Nature* **289**, 24–27 (1981)
4 Ransford, G A, Finnerty, A A & Collerson, K D. *Nature* **289**, 21–24 (1981)
5 Helfenstein, P, Parmentier, E M & Head, J W. *IAU Colloq* No 57 (1980)
6 Crain, I K. in *Random Processes in Geology* (ed Merriam, D F) 3 (Springer, New York, 1976)
7 Lachenbruch, A H *Mechanics of Thermal Contraction Cracks and Ice Wedge Polygons* (Waverly, Baltimore, 1962)
8 Goudsmit, S *Rev Mod. Phys.* 17, 321–322 (1945)
9 Miles, R E *Proc natn Acad Sci U S A* **52**, 1157–1160 (1964), *Math Biosci* **6**, 85–127 (1970), *Sankhya, Ser A* 33, 145–174 (1971)
10. James, A V. G *J Geol* **28**, 458–469 (1920)
11. Howes, L I *Am J Sci* **246**, 138–149 (1948)
12 Nelson, R A. *AAPG Bull.* **63**, 2214–2232 (1979).
13 Beard, C. N *Bull. geol Soc Am* **70**, 379–382 (1969)
14 Peck, D L & Minakami, A *Bull geol Soc Am.* **79**, 1151–1166 (1968)
15 Kindle, E M *J Geol* **25**, 135–144 (1917)
16 Kindle, E M *Am J Sci* **5**, 329–330 (1923)
17. Longwell, C R *Am J Sci* **15**, 135–145 (1928).
18 Lang, W B. *Science* **98**, 583–584 (1943)
19 Neal, J T & Motts, W S. *J Geol* **75**, 511–525 (1967)
20 Badgley, P C *Structural and Tectonic Principles* (Harper and Row, New York, 1965)
21 Christiansen, F. W *Science* **139**, 607–609 (1963)
22 Washburn, *Bull geol Soc Am* **67**, 823–866 (1956)
23 Iranpanah, A & Esfandiari, B *Photogram Engng Remote Sens* **46**, 225–229 (1980)
24 Shoemaker, E. M, Squires, R L & Abrams, M J *NASA Tech Rep 32-1597* (1975)
25 Soderblom, L A *IAU Colloq* No 57, (1980)
26 Schenk, P M & Seyfert, C K. *EOS* **61**, 286 (1980)
27 Parmentier, E. M., Helfenstein, P. & Head, J. W *Abstr 11th Lunar Sci Conf* 860–862 (1980)
28 Melosh, J *Icarus* **31**, 221 (1977)
29 Pieri, D C, Ransford, G A & Collerson, K *IAU Colloq* No 57 (1980)

Europa's petrological thermal history

G. A. Ransford*, A. A. Finnerty* & K. D. Collerson†

* Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103
† Research School of Earth Sciences, Australian National University, Canberra, ACT 2600, Australia

The thermal evolution of Europa is described by a new path of geophysical development. Petrological considerations suggest that heat may have been transported within Europa by sub-solidus convection, and that a large fraction of Europa's water content is retained as hydrated silicates.

We have reconsidered the thermal evolution of Jupiter's second galilean satellite, Europa, from its accumulation from circum-jovian debris until the formation of surface features imaged by Voyager. By including the petrological sequence that probably accompanied Europa's thermal evolution, a path of geophysical development which differs from previous studies[1–3] is indicated. This path gives a present geophysical state that seems to be more in concert with the Voyager observations than earlier models. The petrological considerations imply that heat transport within Europa by sub-solidus convection might have been important throughout much of that satellite's history. We also suggest that a large fraction of the water of Europa is retained in the interior in the form of hydrated silicates.

Thermal history

All previous studies of Europa's thermal history have assumed that the satellite was a mixture of ice and silicates after assembly from the primordial circum-jovian nebula. In these models, Europa is heated by the decay of long-lived radioactive isotopes, and heat is internally redistributed by conduction. High temperatures are reached rapidly, resulting in a differentiation episode. Silicates are segregated into a core of some 1,450 km radius, with a liquid water mantle of 75–100 km thickness, and a 40-km thick ice crust at the surface[3,4]. Cassen *et al.*[5] have studied further evolution from this state within the ice crust and water mantle due to tidal heating effects, and they have concluded that, if once formed, there are conditions in which the water mantle can remain liquid indefinitely. Schubert *et al.*[6] have argued, however, that solid state convection in an ice crust is likely to remove heat from a liquid water mantle sufficiently quickly to cause freezing of the mantle. Cassen *et al.*[7] have improved the earlier calculations, and confirm that persistence of a liquid water mantle is very unlikely.

The total amount of water required to dilute the density of a lunar or Io type body ($\rho = 3.5$) to Europa's value ($\rho \approx 3.0$) is ~5 or 6 wt%, compared with 35 or more wt% for Ganymede and Callisto. Such a small percentage of water need not have started in Europa as ice, because there are many hydrated minerals that

415

Table 1 Weight composition and cation proportions of minerals on Europa

	Weight proportions			Cation proportions (4-oxygen basis)		
	(1)	(2)	(3)	(4)	(5)	(6)
SiO_2	33.1	35.94	35.40	0.947	1.018	1.000
MgO	23.5	25.52	25.65	0.997	1.071	1.083
FeO	35.5	38.54	38.95	0.848	0.911	0.917
Al_2O_3	2.4			0.080		
CaO	2.3			0.070		
Na_2O	1.1			0.061		
NiO	1.9			0.043		
Total	100.0	100.0	100.0	3.043	3.000	3.000

(1) Average composition (wt%) of principal components of Type 1 carbonaceous chondrites on a C-, S- and H_2O- free basis, from ref. 33.

(2) Chondritic composition of column (1), normalized to $SiO_2 + MgO + FeO = 100.0$.

(3) Calculated composition (wt%) of an intermediate olivine with 46 mol% Fe-end member.

(4) Cation proportions for the chondritic composition in column (1).

(5) Cation proportions for (1), normalized to $Si + Mg + Fe = 3$.

(6) Cation proportions for intermediate olivine in column (3).

could be stable in a relatively cool planet of chondritic composition (for example, brucite ~30 wt% H_2O, serpentine ~14 wt% H_2O, chlorite ~12–15 wt% H_2O, talc ~5 wt% H_2O, micas ~2–4 wt% H_2O, amphiboles ~2 wt% H_2O). We consider the possibility here that water was present in Europa as water of hydration from the start.

Water of hydration hypothesis

There are two reasons which indicate that Europa might have started its thermal evolution with most, if not all, of its water in the form of water of hydration. The first is a consideration of the likely pressure–temperature conditions in the Europa zone of the circum-jovian nebula during the condensation of the material that eventually becomes the satellite. Fanale *et al.*[3] reconstructed the circum-jovian nebula in an attempt to understand the simultaneous presence of ice on Ganymede and the 3.0 density of Europa, taking into account the possible earlier luminosity of Jupiter. Whether Io condensed from the circum-jovian cloud without H_2O, or H_2O was subsequently lost from Io, all reconstructions that give high temperatures for the condensation conditions at Io's orbit result in conditions of high humidity and moderate temperatures at Europa's distance. In such conditions, hydrous silicates may condense stably (reviewed in ref. 8).

Accretional heating during the formation of Europa could have dehydrated these condensation products as they accumulated, but most accretion scenarios lead to the opposite conclusion. Harris[9] has shown that the matter accumulated into the galilean satellites had to be predominantly circum-jovian in origin for them to have survived their accretion. Ransford *et al.*[10] discussed how Jupiter completely dominates the dynamics of its circum-jovian swarm of planetesimals, making the problem of accretion in circum-jovian orbit different from accretion in circum-solar orbit. The growing planets exert sizeable perturbations or objects within the circum-solar swarm, which 'pump up' random velocities, and make for relatively high velocity collisions between the forming planets and circum-solar planetesimals. In the circum-jovian accretion case, the forming satellites do not randomize velocities, because of the overwhelming close presence of Jupiter, and relative velocities between circum-jovian planetesimals and the forming satellites are characteristic of the difference in velocity between two objects in nearby orbits. These velocities are much smaller than those in the circum-solar accretion cases, and the attendant heating is expected to be much less. In addition, the accumulation of the small bodies in circum-jovian orbit is thought to have been a moderate time scale process of 10^5–10^6 yr or longer[9,11], which would not raise temperatures inside Europa to above the dehydration temperature for moderate to low temperature accreting material. Thus, the hydrous minerals formed at the condensation stage were probably not dehydrated by the accumulation process.

The second reason for considering water of hydration contrasts with the first. If the circum-jovian nebula considerations cited above are not true, (if Europa does form as a mixture of ice and silicates), subsequent thermal evolution will melt the ice, which will then hydrate the surrounding silicates, forming minerals such as serpentine, brucite and chlorite. These reactions would occur in the low-to-moderate temperature environment that is presumed for Europa's interior early in its thermal evolution provided the appropriate bulk composition is present within the satellite.

If a volatile-free chondritic bulk composition (+5 wt% H_2O) is assumed for Europa, the anhydrous components are equivalent in composition to 92 wt% olivine with 45 wt% (64 mol%) Mg-end member (forsterite) (Table 1). With the presence of 5 wt% H_2O, therefore, reactions in which olivine is hydrated to give serpentine + brucite or chlorite + brucite are most appropriate to Europa. The phase equilibria for two such reactions from Kitahara *et al.*[12] are reproduced in Fig. 1. A curve representing the stability of clinochlore, a chlorite of composition $Mg_5Al_2Si_3O_{10}(OH)_8$ using data from refs 13, 14, is also plotted in Fig. 1. The presence of Fe, and of phases other than clinochlore, probably lowers the dehydration temperature of the chlorite by perhaps up to 200 °C. A curve for the maximum stability of brucite[15] is also shown. For comparison, pressure at the centre of Europa is calculated to be ~31 kbar.

The role of Fe in the stability of serpentine and chlorite is not well known for this bulk composition. Serpentine in terrestrial occurrences generally contain <1.5 wt% Fe_2O_3 or FeO (ref. 16). Chlorite in terrestrial occurrences generally contains some Al, while $Fe^{2+}/Fe^{2+} + Mg$ covers the complete range 0–1. In the temperature range below 600 or 700 °C it is likely that either immiscible Mg-rich and Fe-rich serpentines, Mg-rich serpentine + Fe-rich chlorite, or an intermediate chlorite with no serpentine, will coexist with brucite throughout the pressure range within Europa. Alternatively, the Fe component may crystallize to magnetite or metal, depending on oxygen fugacity. Such high density materials could separate to form a core.

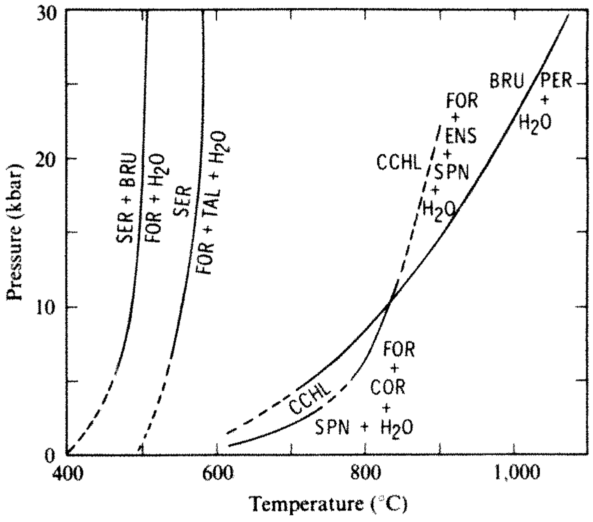


Fig. 1 Stability curves for hydrous minerals and assemblages. Data for forsterite (FOR), serpentine (SER) bulk compositions from ref. 17; clinochlore (CCHL) bulk composition at pressures above 3 kbar from ref. 13; CCHL below 3 kbar from ref. 14; brucite (BRU) composition from ref. 15. Curves are dashed where extrapolated. Other minerals talc (TAL), cordierite (COR), spinel (SPN), enstatite (ENS), periclase (PER).

The products of the hydration reactions represented in Fig. 1 contain 11.7–16.1 wt% H_2O, with 13.7 wt% for the 'chondritic' intermediate olivine composition of Table 1. Using the latter percentage as a likely value for chondritic composition, the entire 5 wt% H_2O in Europa would be consumed if reaction went 36.4% to completion. If water is evenly distributed within Europa, the interior would consist of about two-thirds olivine and one-third serpentine and/or (chlorite+brucite). Alternatively, all the water could be concentrated in an outer hydrated layer composed of serpentine and/or chlorite+brucite, having a thickness of ~270 km.

Discussion

While the water of Europa may once have been incorporated in the interior of the satellite as hydrous silicates, it is not obvious that the water currently exists in the interior after 4,500 Myr of thermal evolution. Using the Moon as an analogue, because its radius is nearly the same as that of Europa, current thermal models employing convective heat transport and materials properties thought to be appropriate for the Earth and the Moon[17–19] predict temperatures well above those at which hydrous minerals dehydrate over most of the interior (Fig. 2). The lithospheres in the models of Fig. 2 are ~300–400 km thick. In a convective Moon model with initial post-accretion temperature of 1,300 °C, Schubert *et al.*[20] calculate a lithospheric thickness of 550 km after 4,500 Myr. While Europa probably started at a much lower initial post-accretion temperature, a planetary object's thermal state after 4,500 Myr has been shown[17,20] to be more dependent on the rheology of the planetary interior than on initial temperature. The upper 200–300 km of the thermal boundary layers or lithospheres in the models cited above are at temperatures below the 500–700 °C dehydration temperatures. Therefore, the hypothesis that most of the water of Europa is incorporated in a relatively cool lithosphere composed of hydrous minerals, which is situated above a hot convecting interior of anhydrous silicates, is consistent with our current knowledge of planetary thermal evolution.

An alternative model may be considered. Tozer[21,22] has suggested that water of hydration affects silicate material properties by decreasing the resistance to creep, that is, by decreasing the effective viscosity. In his convection models for the Moon, the term for temperature dependence of assumed newtonian viscosity was scaled to the solidus temperature of hypothetical rocks Because water lowers the solidus temperature of basalt, the viscosity of the hydrous model Moon is lowered by a factor of ~20, and consequently convection can take place at temperatures below 500 °C. A thermal profile for a Moon composed of "water saturated basalt" from Tozer's work is incorporated in Fig 2. If such a model is applied to Europa, it is clear that hydrated silicates can remain stable throughout the interior, because the low-temperature convection transports heat at a rate sufficient to maintain the interior below the dehydration temperatures.

While "it is generally accepted that subsolidus creep is negligible at temperatures below about 800°C"[20], and therefore that hydrous minerals are unstable in regions that are convecting, this prediction is based on values of materials properties for anhydrous silicates. Hydrolytic weakening of silicate minerals has been convincingly demonstrated only for quartz[23], despite the work of Blacic[24] on olivine. The presence of water weakens the rocks under uniaxial compression, including the serpentinite[25] and chloritite[26] of special interest for our model. However, this weakening is probably a pore-pressure effect[27] and is likely to be ineffective at high pressures (>10 kbar) where pore volume is expected to be negligibly small Additional information on the rheology of 'wet' serpentinite and chloritite will be needed before the possibility of convection within hydrated mineral assemblages can be properly evaluated.

It is expected that some free water gains access to the surface during this evolution, but not in vast quantities. Most of the water present in Europa can be retained in the hydrated minerals in the interior. We predict, therefore, that the ice layer on the surface of Europa is much thinner than the 150 km that had been estimated before the Voyager mission. A thin ice crust on Europa has important implications for tidal coupling and surface morphology (such as cracking—see refs 28, 29 for more details).

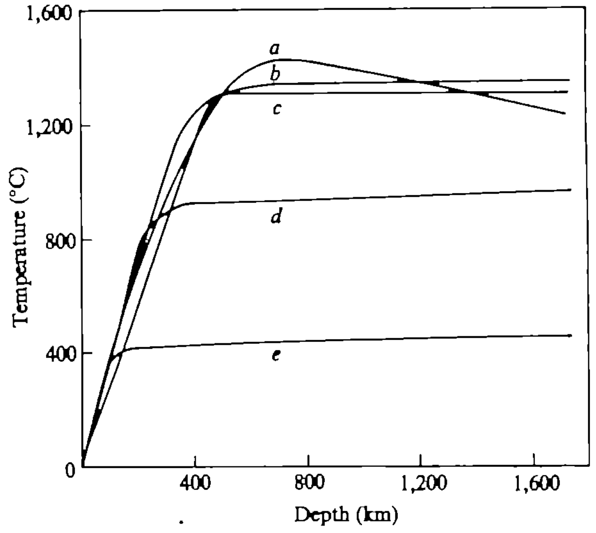


Fig. 2 Spherically averaged temperature–depth curves for the Moon Data for *a*, ref. 34, *b*, ref. 19; *c*, ref 18, *d*, ref 22 (anhydrous dunite model); *e*, ref. 22 (water-saturated basalt model).

The geophysical state that results from a convective evolution agrees with the basic observations about Europa. The synchronous rotation of the satellite would result soon after the subsolidus convection began because of the triaxiality that such patterns could induce in its body[31]. Tidal energy dissipation could involve the entire satellite, as opposed to the crust and core arrangement of the earlier estimates of the thermal history[5]. The completely solid Europa described here would remove a heuristic concern about the earlier Europa—what are the dynamics of a crust with almost no coupling to the core during a tidal interaction? Do they interact with Jupiter independently and librate with respect to each other?

The cracking patterns on Europa's surface suggest some sort of global deformation mechanism[28,32] and internal convective activity would certainly provide such deformation. This aspect will be discussed in more detail by Finnerty *et al.*[29]. The distortions of the silicate surface that are created by the areas of upwelling and subduction in the convective model can fracture the overlying thin layer of ice, generating crack patterns which can be tested against the observed patterns for various orders of convective activity.

Conclusions

We have argued that most of Europa's complement of water can be locked up in the form of hydrated silicates in a thick convective boundary layer or possibly even throughout most of the body of Europa. The water of hydration is in the minerals chlorite and/or serpentine, and brucite, and solid state convection probably keeps a sufficient volume of the planet at temperatures below the dehydration temperatures for these minerals to incorporate all the water needed to explain the density of Europa. With most of the water tied up in hydrated silicates, the ice layer at Europa's surface is probably quite thin, perhaps only a few kilometres.

We thank D. L. Matson, F. P. Fanale, R. J. Phillips, S. Daly, E. Ivins and D. Pieri for helpful discussions. This work represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS 7-100, sponsored by NASA.

Received 5 August; accepted 8 December 1980.

1. Lewis, J. S. *Science* **172,** 1127–1128 (1971).
2. Consolmagno, G. J. & Lewis, J. S. *Icarus* **34,** 280–293 (1978).
3. Fanale, F. P., Johnson, T. V. & Matson, D. L. in *Planetary Satellites* (ed. Burns, J.) 379–405 (1977).
4. Smith, B. A. *et al. Science* **206,** 927–950 (1979).
5. Cassen, P., Reynolds R. T. & Peale, S. J. *Geophys. Res. Lett.* **6,** 731–734 (1979).
6. Schubert, G., Stevenson, D. J. & Ellsworth, K. *Icarus* (submitted).
7. Cassen, P., Peale, S. J. & Reynolds, R. T. *Geophys. Res. Lett.* (submitted).
8. Grossman, L. & Larimer, J. W. *Rev. Geophys. Space Phys.* **12,** 71–101 (1974).
9. Harris, A. W. *Icarus* **34,** 128–145 (1978).
10. Ransford, G. A., Matson, D. L. & Johnson, T. V. *EOS* **60,** 307 (1979).
11. Cameron, A. G. W. & Pollack, J. B. in *Jupiter* (ed. Gehrels, T.) 61–84 (1976).
12. Kitahara, S., Takenouchi, S. & Kennedy, G. C. *Am. J. Sci.* **264,** 223–233 (1966).
13. Fawcett, J. J. & Yoder, H. S. Jr *Am. Miner.* **51,** 353–380 (1962).
14. Chernosky, J. V. Jr *Am. Miner.* **59,** 496–507 (1974).
15. Irving, A. J., Huang, W. L. & Wyllie, P. J. *Am. J. Sci.* **277,** 313–321 (1977).
16. Deer, W. A., Howie, R. A. & Zussman, J. *Rock-forming Minerals* Vol. 3 (Wiley, New York, 1962).
17. Turcotte, D. L., Cooke, F. A. & Willeman, R. J. *Proc. 10th Lunar planet. Sci. Conf.* 2375–2392 (1979).
18. Schubert, G., Young, R. E. & Cassen, P. *Phil. Trans. R. Soc.* A**285,** 523–536 (1977).
19. Cassen, P. *et al. Phys. Earth planet. Inter.* **19,** 183–196 (1979).
20. Schubert, G., Cassen, P. & Young, R. E. *Icarus* **38,** 192–211 (1979).
21. Tozer, D. C. *Phys. Earth planet. Inter.* **6,** 182–197 (1972).
22. Tozer, D. C. *The Moon* **9,** 167–182 (1974).
23. Kirby, S. H. & McCormick, J. W. *Bull. miner.* **102,** 124–137 (1979).
24. Blacic, J. D. *Flow and Fracture of Rocks* (eds Heard, H. C., Borg, I. Y., Carter, N. L. & Raleigh, C. B.) 109–115 (American Geophysical Union Geophysical Monogr. 16, 1972).
25. Raleigh, C. B. & Paterson, M. S. *J. Geophys. Res.* **70,** 3965–3985 (1965).
26. Murrell, S. A. F. & Ismail, I. A. H. *Tectonophysics* **34,** 207–258 (1976).
27. Patterson, M. S. *Experimental Rock Deformation—The Brittle Field* (Springer, New York, 1978).
28. Pieri, D. *Nature* **289,** 17–21 (1981).
29. Finnerty, A. A., Ransford, G. A., Pieri, D. & Collerson, K. D. *Nature* **289,** 24–27 (1981).
30. Runcorn, S. K., *Proc. R. Soc.* A**296,** 270–284 (1967).
31. Roberts, P. H. *Mathematika* **12,** 128–137 (1900).
32. Helfenstein, P & Parmentier, E. M. *Proc. 11th Lunar planet. Sci. Conf.* (submitted).
33. Ringwood, A. E., *Composition and Petrology of the Earth's Mantle*, 192 (McGraw-Hill, New York, 1975).
34. Turcotte, D. L., Hsui, A. T., Torrance, K. E. & Oxburgh, E. R. *J. geophys. Res.* **77,** 6931–6939 (1972).
35. Schubert, G. *Ann. Rev. Earth planet. Sci.* **7,** 289–343 (1979).

Is Europa surface cracking due to thermal evolution?

A. A. Finnerty*, G. A. Ransford*, D. C. Pieri* & K. D. Collerson†

*Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103
† Research School of Earth Sciences, Australian National University, Canberra, ACT 2600, Australia

We propose that some of the surface features of Europa may have originated by processes within a model planet in which hydrated silicates are stable.

THE images of Europa obtained from Voyager display remarkable curvilinear albedo features that range in scale from the limits of resolution to truly global. It has been proposed that most of these features (the lineae) are fractures, and that some (the flexi) are ridges[1–6]. We point out that some surface features of Europa may have originated by processes within a model planet in which hydrated silicates are stable, as proposed by Ransford *et al.*[7]. Because we deal with only a few types of the features that have been observed, additional processes, or perhaps entirely different processes than proposed here, may be operating. Because of the paucity of data bearing on remote planetary objects, however, our understanding of these processes can be advanced only by constructing successively more realistic models that can be tested against the limited observations.

In the Ransford *et al.*[7] model for Europa, a region of convecting anhydrous silicates is overlain by a lithosphere, the upper 200–300 km of which is composed of hydrous minerals. The ice crust need not be thick (not nearly the 75–150 km of ice and possibly water in pre-Voyager models) because all the H_2O needed to explain the lower density of Europa compared to Io can be contained within the hydrated silicates. To relate surface features of Europa to this model, we first examine the mechanical properties of hydrated silicates and ice.

Mechanical properties

The compressive strengths and ductilities of serpentinite, of partially serpentinized peridotite, and of chloritite have been measured at confining pressures up to 6.6 kbar at various temperatures [8,9]. Complex but consistent behaviour was observed. At temperatures below the dehydration temperatures (500–700 °C) serpentinites and chloritites are about as strong as granite, and brittle, becoming more ductile as confining pressure and temperature increase. As temperatures approach the dehydration temperatures, the hydrous rocks weaken slightly, but at the dehydration temperatures, strengths decline drastically to <500–1,000 bar and the specimens become much more brittle. This is apparently due to release of free water, because at temperatures below the dehydration temperatures, the introduction of free water causes similar weakening and embrittlement. This behaviour has been related to some types of faulting on Earth[8–10]. If free water is present in the lithosphere of Europa, therefore, fracturing should occur at quite low stress levels.

The mechanical behaviour of water ice is very similar to that of anhydrous silicates. At a small fraction of its melting temperature it deforms in a brittle fashion[11], while at a high fraction of melting temperature it deforms plastically by a creep mechanism (reviewed in ref. 12). The mechanical behaviours of thin and thick ice crusts on Europa should consequently differ. (We do not consider a liquid water mantle for Europa because new calculations[13] suggest that the previously proposed liquid zone[14] is very unlikely.)

Thickness of the ice crust

Using a near-surface (below diurnal influences) temperature for Europa of 93 K (ref. 15), and a thermal gradient of $3\ \mathrm{K\ km^{-1}}$ (from the conductive region of the thermal models discussed in ref. 7), a temperature of 60% of the melting temperature of ice will be reached at ~24 km depth. If such a relative temperature is taken as marking the brittle-to-ductile transition, then any ice existing below 24 km depth will deform by plastic rather than brittle mechanisms. While such deep ice will be capable of transmitting isotropic stresses and allow shallower brittle ice to fracture, any nonisotropic stress from below will be compensated by lateral flow within warm ice. The presence of a reticulate or polygonal fracture pattern in the anti-jovian region of Europa[6] on a scale larger than the 150 km maximum depth of an ice layer, but on a smaller than planetary scale, suggests that nonisotropic stresses have propagated through the ice crust and that the crust is consequently thinner than about 24 km, at least in the anti-jovian region. Smith *et al.*[1] have also suggested that the ice layer may be thin (<50 km) based on the implications of two terrain types and very low surface relief. Therefore, we will consider the ice crust to be thin and brittle, acting as a recorder of stresses imposed on the underlying hydrated silicate mantle.

0028-0836/81/020024—07$01.00

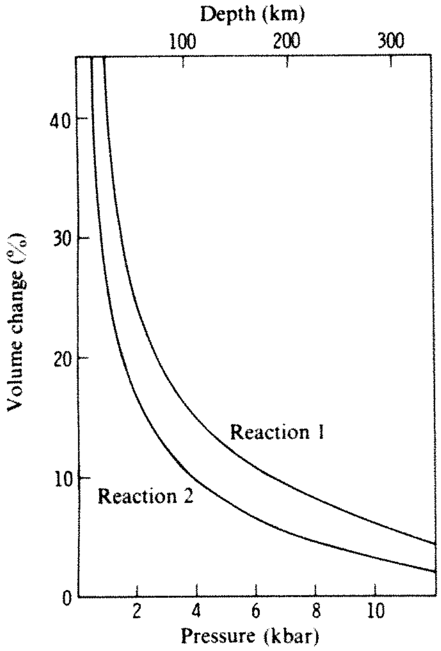


Fig. 1 Relative volume changes as a function of depth for dehydration reactions involving serpentine.

Models for two sets of surface features on Europa have been developed: (1) the long, relatively dark lineaments that lie on great and small circles, and frequently possess bright medial stripes (principally the Types 1 and 3 lineaments of ref. 6); (2) the complex reticulate network of fractures in the anti-jovian region, surrounded by the series of cuspate ridges defining a small circle about 60° outside the anti-jovian point (Types 5 and 6 lineaments of ref. 6).

Global scale lineaments

The great lengths (>1,000 km in many cases) and distribution (for example, Asterius Linea defines a nearly complete great circle) require a global scale mechanism for formation of Types 1 and 3 lineaments. A possible mechanism is planetary expansion. Volume changes due to thermal expansivity can be positive or negative depending on the thermal history (see ref. 16). Dehydration of hydrated silicates at relatively low pressures, however, results in large positive volume changes, and may be responsible for formation of Types 1 and 3 lineaments.

Dehydration may be represented by reactions such as:

$$\underset{\text{serpentine}}{Mg_3Si_2O_5(OH)_4} + \underset{\text{brucite}}{Mg(OH)_2} \rightleftharpoons \underset{\text{forsterite}}{2Mg_2SiO_4} + \underset{\text{fluid}}{3H_2O} \qquad (1)$$

and

$$\underset{\text{serpentine}}{5Mg_3Si_2O_5(OH)_4} \rightleftharpoons \underset{\text{forsterite}}{6Mg_2SiO_4} + \underset{\text{talc}}{Mg_3Si_4O_{10}(OH)_2} + \underset{\text{fluid}}{9H_2O} \qquad (2)$$

Other reactions involving iron-bearing components and/or chlorite lead to similar conclusions. Considering a closed system in which the total H_2O cannot vary, the change in volume for reactions (1) and (2) may be calculated from the data summarized in Table 1. The molar volumes of the minerals do not change dramatically with temperature and pressure, and in any case these changes approximately cancel in calculating the change in volume, so their room temperature, one atmosphere volumes are used. For H_2O, however, the molar volume does change drastically with temperature and pressure. The dehydration temperature of serpentine is ~500 °C, and is nearly independent of pressure. Therefore, volume changes for reactions (1) and (2) are calculated as a function of pressure at 500 °C and plotted in Fig. 1. A depth scale for Europa is also illustrated. Molar volume data for H_2O at pressures >10 kbar are not available, but the volume probably does not decrease below 14.4 $cm^3\,mol^{-1}$ at 30 kbar and 500 °C. This corresponds to volume changes at the centre of Europa that are close to zero.

If Europa accreted at low temperatures[7], hydrated minerals would initially be distributed throughout the satellite. The temperature at a given depth would increase with time due to decay of radioactive isotopes. Assuming an adiabatic or superadiabatic temperature gradient, the hottest location is at the centre of the planet, regardless of whether solid state convection commences. At some time the temperature at the centre of the planet would rise above 500–700 °C, and dehydration would commence. A dehydration front would coincide with the upward evolution of the 500–700 °C isotherm. If H_2O did not escape upwards, the assemblage of anhydrous minerals + vapour would expand relative to the original hydrated assemblage.

If H_2O can percolate upwards, it will react with anhydrous minerals in the layer above the 500–700 °C isotherm to increase the hydrous mineral content from about 40% to as high as 100%. Any excess H_2O will migrate still higher and hydrate additional layers. In this case, the dehydrated region decreases in volume while the layer of increased hydration expands. Because ΔV is greater with shallower depth, however, the net volume change is positive.

This expansion creates stress in the overlying hydrated silicates, and at the same time the dehydration process provides the free water that completes the hydration of these rocks. Eventually, the water is concentrated in a completely hydrated layer extending from the surface to ~270 km depth[7]. Additional water percolating from below can no longer be bound in hydrated minerals, and thus weakens and embrittles these rocks even though they are below their dehydration temperatures. The hydrated silicate shell may be modelled as a spherical pressure vessel, in which tangential and radial stresses due to dehydration are at maximum at the inner surface and minimum at the outer surface. For a fracture to propagate to the surface, the crack probably needs to be filled with water, provided by the dehydration, as it grows. As Anderson and Grew[17] noted, water is particularly effective at promoting crack growth in silicates by stress corrosion.

The presence of bright medial stripes within lower albedo lineaments hundreds of kilometres long and tens of kilometres wide must be explained by any hypothesis of the origin of these features. Malin[2] noted that the bright 'cores' are in places hundreds of metres above the surrounding terrain. Also, we note that whereas the white medial stripe is usually very uniform in width within a particular lineament, the darker portion is

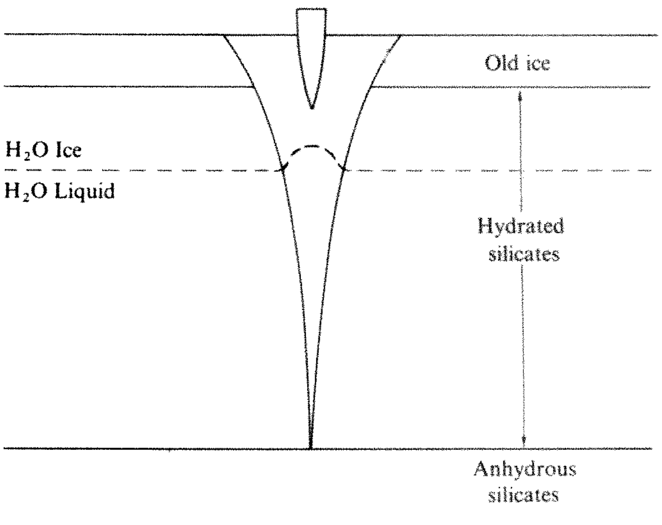


Fig. 2 A representation in cross-section of the formation of planetary scale dark lineaments with bright medial stripes. The shaded region is xenolith-rich H_2O 'magma'. The dashed line represents the freezing temperature of ice. The wedge at the top is new ice extruded upwards as the freezing isotherm migrates downward during thermal re-equilibrium of the magmatic column.

Table 1 Molar volumes of components ($cm^3\ mol^{-1}$)*

Serpentine	$Mg_3Si_2O_5(OH)_4$	108.5
Brucite	$Mg(OH)_2$	24.63
Forsterite	Mg_2SiO_4	43.786
Talc	$Mg_3Si_4O_{10}(OH)_2$	136.25
Vapour	H_2O	18.169†

* Mineral data from Robie *et al.*[26].

† Data from Burnham *et al.*[27] for H_2O vapour at 500 °C, 10 kbar.

quite variable and, where it widens on one side of the medial stripe, there is frequently a corresponding widening on the opposite side.

We suggest the following explanation (Fig. 2), analogous perhaps to CO_2-driven kimberlite eruptions on Earth. When a crack is initiated at the base of the lithosphere, its propagation upwards is aided by stress corrosion as water from the dehydrating region fills the crack. Because this aqueous fluid is highly corrosive to Si–O–Si bonds and of low viscosity, the crack tip can propagate rapidly[17]. Relatively high fluid velocities may be attained, in which case xenoliths can be broken from the sides of the crack and incorporated into the ascending fluid. As the surface is approached, velocity of the fluid will diminish due to loading in the 'magmatic' column and loss of heat to the walls of the conduit. A breccia of frozen water and xenolithic silicates and ice, of lower albedo than the surrounding icy crust, will fill the upper reaches of the conduit, while in the lower portion H_2O will remain liquid and xenoliths can settle out. As the crack propagates laterally, presumably following a great circle, eruptions of this brecciated material with varying vigour will cause variations in the surface width of the low albedo material. At the conclusion of this process, the isotherm coincident with the water-ice transformation will be bowed upwards within the 'breccia pipe'. As this isotherm evolves downwards within the fracture towards thermal equilibrium, additional water will freeze, causing expansion and a new stress. Because the depth of ice breccia in the crack will be least at the centre of the crack, a second stage fracture will form within the ice breccia and in the middle of the initial crack. New ice will extrude along the new fracture, rising above the surrounding terrain due to the liquid-ice volume change. Inasmuch as the new ice is essentially xenolith-free, it will have higher albedo than the ice breccia on either side.

Features about the anti-jovian region

A concentration of fractures interpreted as tensile features, in polygonal or reticulate patterns has been noted[4,6,18] over a region near the anti-jovian point on Europa. This interpretation, based on an analysis of the frequency distribution of polygon sides, has been justified by Pieri[6]. The polygonal pattern is formed by the intersections of Types 1 and 2 lineaments and is subdivided to the limits of resolution by the Type 5 lineaments of Pieri. For this modelling exercise, we also assume that these are fractures formed in a tensional environment.

Because of their geometrical relationship to the tensional fractures in the anti-jovian region, we consider the flexi, or Type 6 lineaments of Pieri in any model of processes occurring in this locality. These cuspate ridges appear to surround the tensile fractures in a ring about 60° out from the anti-jovian point, and have up to several hundred metres of relief. Their maximum extent is not known because of incomplete high-resolution coverage and unfavourable lighting angles. There is no recognized evidence of erosional or magmatic processes on Europa to form these ridges. If they are of tectonic origin, they probably formed within a region of compressional stress. Strom *et al.*[19] described lobate scarps and ridges up to 500 km long and 3 km high on Mercury, uniformly distributed over the hemisphere imaged by Mariner 10, and attributed their formation to crustal compression during thermal contraction of Mercury. While the distribution of Europan cuspate ridges differs from that on Mercury, and we do not yet know the profiles of the ridges, we feel that the interpretation as compressional tectonic features is most likely. In the absence of data to the contrary, we assume that the compressional features formed contemporaneously with the tensional features discussed above.

We are faced, then, with a broad region of extension centred near the anti-jovian point, surrounded by a region of compression. Based on the radial distribution of tensional features, doming, and concentration of volcanism in the Tharsis area of Mars, Carr[20] suggested that a region of upwelling within a convecting mantle was centred under Tharsis. Solomon and Chaiken[16] rejected this interpretation because no regions of compression corresponding to zones of return flow are apparent on Mars. Although this model may not work for Mars, for Europa, we suggest that a region of upwelling in the convecting anhydrous silicate interior, surrounded by a region of return flow, may have been located beneath the hydrated silicate mantle at roughly the anti-jovian point during the formation of the tensile and compressional features. The lineaments formed above the convection cell by brittle deformation within the hydrated silicate lithosphere.

We still need to explain the location of the convective upwelling at the anti-jovian region. This is probably not accidental, because the redistribution of hot, low density and cold, high density material by convective processes alters the moments of inertia of the satellite, and can thus influence the orientation of the satellite when it achieves orbital period–spin rate lockup with the primary.

In combination with the mass distribution associated with equatorial bulging due to rotation, a second-order convection pattern would cause the triaxiality needed for the tidal locking of the rotation period of Europa. This process has been discussed extensively in connection with heat transfer and nonhydrostatic shape of the Earth's Moon (see refs 21–23), and results in diametrically opposed zones of upwelling becoming aligned with the radial to the primary. In the case of the Moon, the maximum shear stress imposed by the nonhydrostatic figure is probably insufficient to fracture the lithosphere[24], and a dynamic process such as convection is not required. For Europa, the 'wet' hydrated silicate lithosphere should be far weaker than the probably anhydrous lithosphere of the Moon, and consequently a dynamic process such as convection should be revealed by failure of the lithosphere.

A second-order convection pattern requires a small core, as might have been provided in Europa by separation of iron or magnetite on deserpentinization[7]. Such a pattern is shown in Fig. 3. The distribution of extensional and compressional fea-

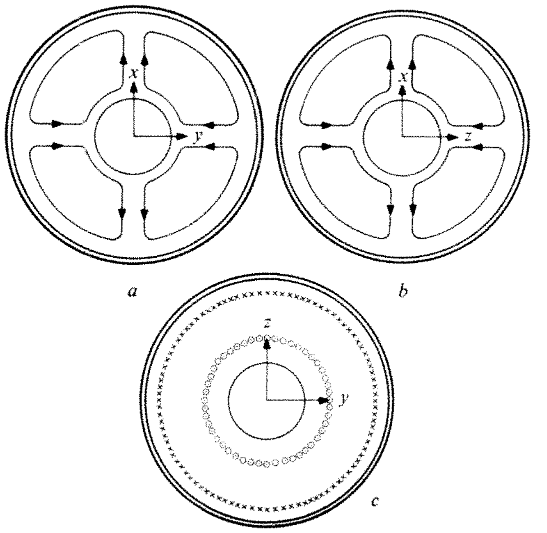


Fig. 3 Idealized second-order convection patterns in three mutually orthogonal projections. Axis of rotation is the z-axis, direction towards Jupiter is along the x-axis, and the y-axis points in the instantaneous direction of motion of Europa in its orbit around Jupiter. ○, Upwelling; ×, downflow.

tures in the anti-jovian region of Europa is remarkably similar to the locations of zones of upwelling and return flow in Fig. 3*c*. However, the effects of other processes may be superimposed on the surface of Europa, and convection within a planetary interior is probably more complicated than second order. Nonetheless, a major second-order component may be present.

Discussion

A spherical pressure vessel model for Europa predicts no preferred orientation for the planetary-scale fractures. Helfenstein and Parmentier[18] have noted, however, a tendency for these lineaments to intersect with an angle of 60–70°. They use the elastic shell theory developed by Melosh[25] to show that the patterns of these lineaments and the location of a region of tension in the anti-jovian region are consistent with relaxation of a triaxial figure composed of two biaxial distortions: an equatorial bulge and a bulge along the radial to Jupiter. They propose a mechanism in which the equatorial bulge relaxes as Europa despins due to tidal interactions, and the anti-jovian bulge relaxes as the orbit of Europa recedes from Jupiter. We are not convinced that such a preferred orientation of global scale fractures exists on Europa because of limited photographic coverage, uncertainty about which fractures were studied by Helfenstein and Parmentier, and the probability that fractures formed by several different processes may have been included in their analysis. There are also two difficulties with their model: the magnitude of stresses imposed by relaxation of the figure of Europa (tens of bars) probably does not exceed the strength of the lithosphere (Banerdt, personal communication); and the tidal bulge cannot have been located in the anti-jovian region until after Europa had despun into a tidally-locked configuration. If preferred orientations of global scale fractures are documented, we suggest that global expansion due to dehydration of the interior provided stresses great enough to fracture the lithosphere, and that superimposed stresses due to tidal despinning may have influenced the orientation of these fractures. The triaxial contribution to the fracture patterns may have been provided by dynamic loading of the lithosphere by convection in the interior.

We re-emphasize the hypothetical nature of our arguments. We hope that these speculations will stimulate new observations and will lead to new models that will improve or supersede the model proposed here.

We thank D. L. Matson, E. Ivins, S. Daly, B. Banerdt, F. P. Fanale and R. J. Phillips for helpful discussions; and D. L. Matson, F. P. Fanale for helpful comments. This research represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS 7-100, sponsored by NASA.

Received 5 August; accepted 8 December 1980.

1. Smith, B. A. *et al. Science* **206,** 927–950 (1979).
2. Malin, M. C. *IAU Colloq.* No. 57, abstr. 7–2 (1980).
3. Helfenstein, P., Parmentier, E. M. & Head, J. W. *IAU Colloq.* No. 57, abstr. 7-4 (1980).
4. Parmentier, E. M., Helfenstein, P. & Head, J. W. *Abstr. 11th Lunar Planet. Sci. Conf.* 860–862 (1980).
5. Pieri, D., Ransford, G. & Collerson, K. *IAU Colloq.* No. 57, abstr. 7-3 (1980).
6. Pieri, D. *Nature* **289,** 17–21 (1981).
7. Ransford, G. A., Finnerty, A. A. & Collerson, K. D. *Nature* **289,** 21–24 (1981).
8. Raleigh, C. B. & Paterson, M. S. *J. geophys. Res.* **70**, 3965–3985 (1965).
9. Murrell, S. A. F. & Ismail, I. A. H. *Tectonophysics* **31,** 207–258 (1976).
10. Heard, H. C. & Rubey, W. W. *Bull. geol. Soc. Am.* **77,** 741–760 (1966).
11. Parameswaran, V. R. & Jones, S. J. *J. Glaciol.* **14,** 305–315 (1975).
12. Weertman, J. in *Physics and Chemistry of Ice* (eds Walley, E., Jones, S. J. & Gold, L. W.) 320–337 (Royal Society of Canada, Ottawa, 1973).
13. Cassen, P., Peale, S. J. & Reynolds, R. T. *Geophys. Res. Lett.* (submitted).
14. Cassen, P., Reynolds, R. T. & Peale, S. J. *Geophys. Res. Lett.* **6,** 731–734 (1979).
15. Gaffney, E. S. & Matson, D. L. *Icarus* (submitted).
16. Solomon, S. C. & Chaiken, J. *Proc. 7th Lunar Sci. Conf.* 3229–3243 (1976).
17. Anderson, O. L. & Grew, P. C. *Rev. Geophys. Space Phys.* **15,** 77–104 (1977).
18. Helfenstein, P. & Parmentier, E. M. *Proc. 11th Lunar planet. Sci. Conf.* (submitted).
19. Strom, R. G., Trask, N. J. & Guest, J. E. *J. geophys. Res.* **80,** 2478–2507 (1975).
20. Carr, M. H. *J. geophys. Res.* **79,** 3943–3949 (1974).
21. Roberts, P. H. *Mathematika* **12,** 128–137 (1965).
22. Runcorn, S. K. *Proc. R. Soc.* **A269,** 270–284 (1967).
23. Cassen, P., Young, R. E. & Schubert, G. *Geophys. Res. Lett.* **5**, 294–296 (1978).
24. Phillips, R. J. & Ivins, E. R. *Phys. Earth planet. Inter.* **19,** 107–148 (1979).
25. Melosh, H. J. *Icarus* **31,** 221–243 (1977).
26. Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. U.S. geol. Surv.* **1452** (1978).
27. Burnham, C. W., Holloway, J. R. & Davis, N. F. *Geol. Soc. Am. Spec. Pap.* **132** (1969).

Sex ratio evolution in a variable environment

E. L. Charnov*, R. L. Los-den Hartogh†, W. T. Jones‡ & J. van den Assem†

* Department of Biology, University of Utah, Salt Lake City, Utah 84112
† Zoölogisch Laboratorium, Rijks Universiteit Leiden, Leiden, The Netherlands
‡ Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

We develop a natural selection model for sex ratio control in a spatially variable environment. Predictions of sex ratio alteration as a function of environmental change are tested in laboratory experiments with two parasitic wasps. Field data from a variety of other organisms also support the model. Finally, we discuss possibilities and difficulties for testing this type of evolutionary model.

SEX allocation theory deals with the impact of natural selection on the allocation of resources to male compared with female reproductive function[1]. For dioecious species the problem of interest is the evolution of the sex ratio; for a simultaneously hermaphroditic species, it is the allocation to sperm versus eggs in each breeding season[2–5]; for a sequential or sex-reversing hermaphrodite, the questions are the sex order (male or female first?) and the proportion of an individual's lifetime devoted to each sex[6–12]. The theory also gives the conditions for a sexual state to be evolutionarily stable. For example, it specifies when some form of hermaphroditism is favoured over dioecy[1–16].

It was previously[1] shown that many sex allocation problems have the property that the equilibrium or ESS[17] condition satisfies a particular optimality principle. The allocation to male versus female function, which is favoured by natural selection, can often be shown to maximize the fitness gains through male function multiplied by the gains through female function[1]. From this product formalism, one can derive the classic results of sex ratio theory. Examples are the 'equal resource into each sex' theorem of Fisher[18], the 'local mate competition' results of Hamilton[19] and several results from the Trivers–Hare theory for eusocial Hymenoptera[20,21]. Here, we shall apply the formalism to sex ratio evolution in a variable environment and will provide a set of laboratory tests for the resulting predictions. As the organisms which stimulated our modelling efforts were parasitic wasps, we shall develop the model with reference to their life histories.

Parasitic wasps in a variable environment

Imagine an outcrossing parasitic wasp which has discrete generations and where the females attack hosts over a wide range of sizes[22–34]. A single egg is laid on each host. The host is paralysed or killed by a sting from the female wasp, so that the total food for her offspring's development is contained in the host at the moment of attack. If the host is small, the wasp larva will have relatively little food and will emerge as a small adult. If

0028-0836/81/020027—07$01.00

the host is large, the resulting wasp will be large. Suppose further that the reproductive consequences of being a large compared with a small adult wasp differ, depending on whether the individual is a male or a female. For example, over her lifetime, a large female may lay 10 times as many eggs as a small female; however, a large male may be only three times as effective as a small male at inseminating females. As the female wasps lay eggs in a variety of host sizes, the environment is thus variable with respect to opportunities for production of sons versus daughters. The mother is assumed to control her sex ratio (proportion male eggs) as a function of host size. Haplodiploid sex determination provides her with a physiological mechanism for this control. Our interest is in predicting what sex ratio she will produce in a given host size. To do this, we first specify the host size distribution and the expected reproductive success for an adult wasp of a given sex and size.

Let $f(x)$ be the probability density function for hosts of size x which are attacked by the female. For simplicity, assume x to be continuous. Let $W_1(x)$ be the relative fitness of a son derived from a host of size x. A male's fitness is measured relative to other males[1], so that $W_1(x)$ scales the ability of a male of size x to inseminate females over his lifetime. Let $W_2(x)$ be the lifetime egg production of a female derived from a host of size x. Both these fitness measures include survivorship to adulthood, which may vary with host size[22]. Finally, let $r(x)$ be the proportion of unfertilized or male eggs laid in host size x. $r(x)$ is controlled by the mother wasp. Using techniques from population genetics, one can show that the equilibrium or ESS[17] $r(x)$ satisfies the following product relationship[1]. It is the function, $r(x)$, which maximizes the product of the fitness gained through the decisions to produce sons and that gained through the decisions to produce daughters, or:
Maximize

$$\left[\int_x f(x)W_1(x)r(x)\,\mathrm{d}x\right]\times\left[\int_x f(x)W_2(x)(1-r(x))\,\mathrm{d}x\right]0\leq r(x)\leq 1 \qquad (1)$$

If we impose the condition that a large daughter gains more in relative fitness than a large son [$W_2(x)/W_1(x)$ increases with x], the solution to equation (1) reduces to a very simple form[1,35]. Only sons should be produced in small hosts and only daughters in large ones. There is a certain host size (τ) where the switch from sons to daughters occurs (Fig. 1*a*). Equation (1) now takes the form:
Maximize

$$\left[\int_0^{\tau} f(x)W_1(x)\,\mathrm{d}x\right]\times\left[\int_{\tau}^{\infty} f(x)W_2(x)\,\mathrm{d}x\right] \qquad (2)$$

Although the model predicts a threshold (τ) for the sex ratio change, it is unlikely that biological data would show such an abrupt transition. A more likely pattern would be that of all sons in small hosts and daughters in large, with a gradual sex ratio transition in between (illustrated in Fig. 1*a*)[1,35]. We expect a gradual transition for two reasons. First, the model assumes time constant fitness and host distribution values. If these vary through time, then so does τ. The quality of a host may also depend somewhat on factors which cannot be known to the mother when she deposits an egg. Selection would then favour a more gradual shift in sex ratio. However, a second reason is that even if the mother wasp made perfect decisions, the experimenter may not be able to classify the hosts perfectly with respect to criteria used by the wasps. For example, differentiation of equation (2) with respect to τ shows that the hosts are theoretically to be classified by the ratio $W_2(x)/W_1(x)$. Those with low ratios are son hosts, those with large are daughter hosts. Let k be the threshold ratio. Now, suppose that the wasp makes perfect decisions, with all hosts above k being given to daughters. If the experimenter, however, uses a measure of host size ($\hat{x}$) which is imperfectly correlated with the $W_2(x)/W_1(x)$ ratio, we will not see the threshold. The experiment would show a gradual shift in sex ratio as a function of $\hat{x}$, our measure of host quality.

In addition to the sex ratio shift with host size, the model makes one other general type of prediction. The concepts of large and small host as discussed above have no absolute meaning. That is, a host is only large or small relative to the other hosts being attacked. To calculate theoretically whether a host is to be a son or daughter host, we must first know the host size distribution which is being attacked. We expect to find an increasing fraction of daughters in larger hosts; however, what constitutes larger may vary from place to place, or time to time. Consider the host marked z on Fig. 1*b*. If the distribution of host sizes is i, z is a large host. If the distribution is ii, z is a relatively small host. Host z should contain a greater proportion of daughters when large (case i) than when small (case ii).

Two other factors implicit in the model deserve further comment. First, the sex ratio under discussion is the ratio at the egg stage. Data often show that survivorship declines in small hosts[22]. If this decline is different for the two sexes, the differential mortality as a function of host size would itself generate a sex ratio shift among emerging wasps, independently of a true sex ratio shift at the egg stage. However, such differential mortality would also contribute to natural selection favouring a shift by the mother wasp. Some evidence indicates that daughters may sometimes be unable to reach pupation size in very small hosts (a pupal parasite of Lepidoptera[34]), but even here a strong sex ratio shift at the egg stage with host size could be demonstrated. The relativity prediction discussed previously is a good control for

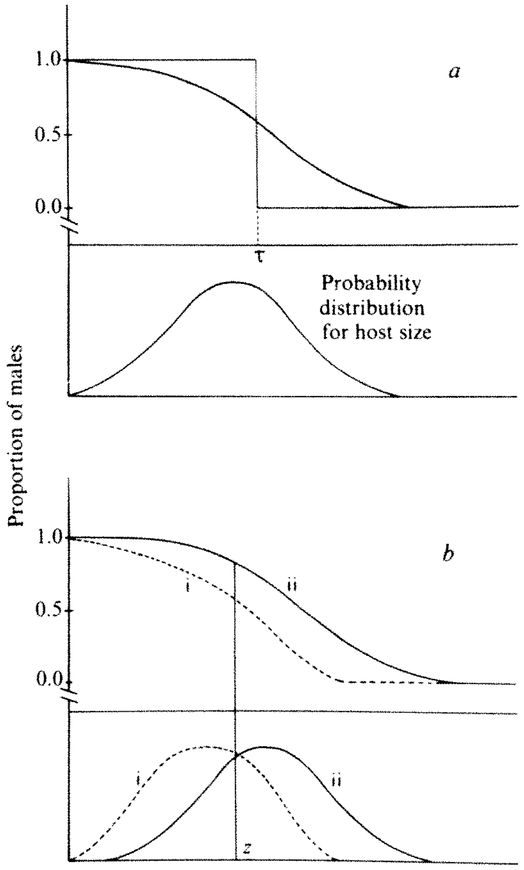


Fig. 1 Sex ratio as a function of host size. *a*, Provided females gain more fitness by being larger than do males (that is, $W_2(x)/W_1(x)$ increases with x), the sex ratio is predicted (equation (2)) to be female biased in large hosts, male biased in small. There is a threshold host size (τ) where a change-over occurs. For reasons discussed in the text, the expected pattern is not a threshold, but more gradual sex ratio shift. *b*, Whether a given host is to be a son host or a daughter host depends on the entire host size distribution. Host z is a relatively large host if the host distribution is i; it is a small host if the distribution is ii. It should contain more males when it is small. Host size distributions are simply illustrative.

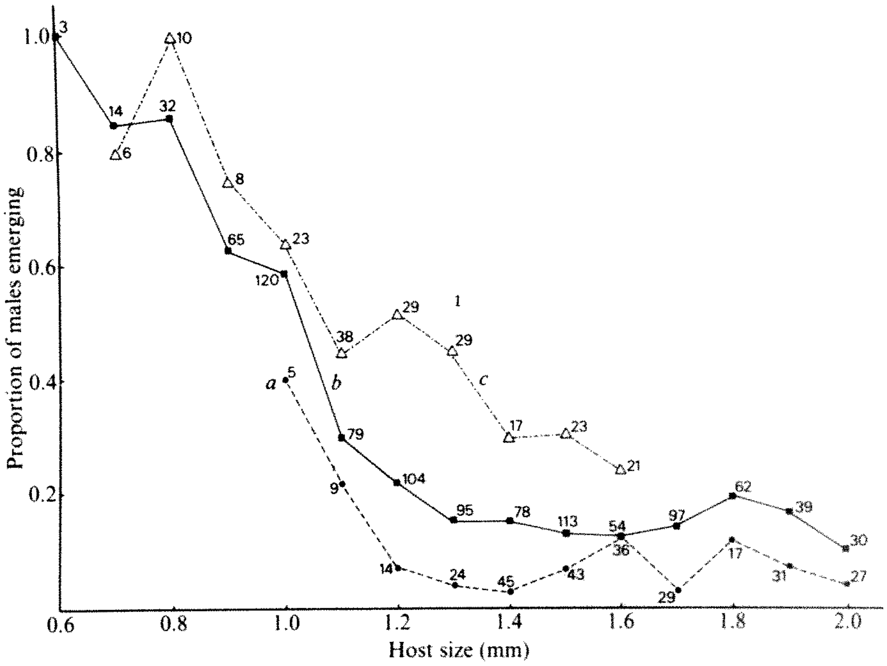


Fig. 2 Sex ratio in *Lariophagus*. The proportion of males emerging is plotted against the size of the hosts (weevil larvae) (tunnel width) which were offered in a 'carousel experiment'. Curve *b* results when female wasps are presented sequentially with 20 hosts of a single size. Curves *a* and *c* result when wasps are presented with an alternating sequence of two host sizes, where the hosts are 0.4 mm different is size. Curve *c* is where the host of interest was the smaller of the two. For example, hosts of 1.4 mm were offered alone and gave a sex ratio of 15% males (curve *b*). When they were offered alternately with 1.8-mm hosts, they gave a sex ratio of 30% males (curve *c*). When offered with 1.0-mm hosts, their sex ratio dropped to 2% males (curve *a*). The number of emerging wasps is given for each data point. The experimental technique is as follows: the weevil *Sitophilus* is cultured on wheat grains at 25 °C. Twice a week freshly emerged weevils are placed on fresh (unfrozen) grain for oviposition and removed again 7 days later. Grains with the necessary age range of larvae can then be irradiated with X rays to measure the exact size of the larvae by measuring tunnel diameter[37]. The grains are then stored at 10 °C for up to 14 days. At 10 °C the larvae do not grow and only a small percentage die. Grains with larvae of a desired size are picked out and used in carousel experiments. The oviposition carousel (described in ref. 30) makes it possible to offer individual female wasps grains with known contents for a specified period of time in a specified order. Inseminated females are isolated from cultures on the day of their emergence and stored for 24 h at 18–20 °C. They are used in a carousel experiment, with a series of 20 grains. Each grain is offered to a female for 2.5 h, then labelled and stored at 25 °C until a weevil or a wasp emerges. If there is no emergence, the grain is opened to see what happened to the weevil larvae. The head width was measured for all wasps.

this effect, as the shift is predicted within a single host size. The model also predicts that the overall sex ratio should be biased towards males. However, this is again sex ratio at the egg stage. If survivorship declines with host size, then at emergence the overall sex ratio may be female biased (because males are put in small hosts).

The second factor is the assumption that the species is outbred. Hamilton[19,36] has developed a theory for certain forms of inbreeding, or nonrandom competition among males for mates, which predicts a generally female-biased sex ratio. In the most extreme cases, all females in a brood are inseminated by their brothers. Here, theory predicts (and data strongly support) that a mother wasp should produce mostly daughters. This 'local mate competition' (LMC) is thus a factor which may alter the overall sex ratio in the direction of females. Although we cannot eliminate it as a factor in our wasp systems, it seems unlikely as a general explanation for a host shift in sex ratio. Long-term (~20 yr or 200 generations) laboratory cultures of one species studied here[22] consist of hundreds of wasps mixed together. There is much mixing of the offspring of individual females. Such conditions are quite the opposite of those postulated by Hamilton and would select against LMC effects. In this same species (*Lariophagus*), other experiments presented the wasps with groups of hosts and sometimes put two or more female wasps together with a batch of hosts. The present experiments presented a single female with one host at a time and they gave essentially the same results as the former. With a strong LMC model, the wasps would be predicted to behave differently in the two conditions. For these reasons, we have derived the predictions from a model with outbreeding (Fig. 1).

On testing the model

A general test of the model would consist of showing that sex ratio varied as a function of host size, and that the daughter/son fitness comparison also changed with host size. Both would have to change in the predicted directions. We would also have to show that differential mortality or LMC could not account for the results. The predicted relative nature of the sex ratio shift itself provides a powerful, qualitative test of the hypothesis. There are at least three ways in which we might test this prediction.

(1) Geographical variation. If in different places the wasp species is associated with different host size distributions, this natural variation can be turned into a test. For instance, if the same host is present in both locations, but is a small host in one place and is large in the other, a sex ratio shift is expected in that host.

(2) Laboratory selection experiments. It should be possible to mimic the effects of geographical variation in the laboratory. If the parasitoids can be reared on only small hosts, or only large hosts, sex ratio shifts through time (over several generations) are expected, if the sex ratio is a heritable character. The small hosts should shift from producing mostly males towards a more equal sex ratio. The large hosts should shift from mostly females towards a more equal ratio.

(3) Temporal variation. If the wasps naturally confront host size distributions which change from generation to generation, they might have the ability to alter their sex ratio on a short time scale, in response to the prevailing distribution. The most reasonable source for a shift in host size utilization would be a shift in host abundance. Because small hosts are poorer for offspring, a mother might be expected to ignore them when confronted with an abundance of larger hosts. This behavioural shift would require the wasps to pick up, store and use information about the distribution. However, if the natural host size distribution was either constant through time, or continued to change within a wasp generation, we might expect the wasps not to respond to short-term shifts in the distribution. We would then require geographical variation or a laboratory selection experiment to test the relativity idea.

We report here the results of a series of experiments on this host size model. Our experiments relate to the abilities of two parasitoid species to alter their sex ratio decisions on a short time scale. We tested the basic model by looking at sex ratio as a function of host size. We then confronted wasps with altered

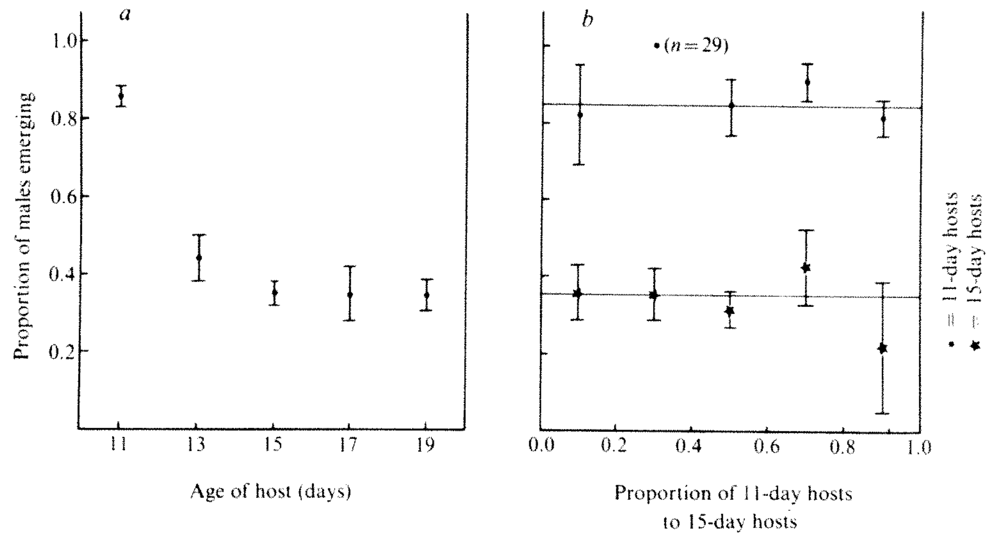


Fig. 3 Sex ratio as a function of host age for *Heterospilis* attacking *Callosobruchus*. *a*, Wasps presented with a single-aged host for a period of 24 h (30°C and 70% relative humidity, host density at ~20 hosts per female wasp). Data given are means ±2 s.e. *b*, Wasps were in an environment consisting of 11- and 15-day hosts presented simultaneously. In the five experiments, the proportion of 11-day hosts varied from 0.1 to 0.9. Data are the mean sex ratios within each host age (±2 s.e.). The lines in *b* refer to the proportion of males found in *a*.

host size distributions and measured the sex ratio shift. Our data indicate that large wasps come from large hosts and that both males and females gain increased reproductive success by being large.

Sex ratio in *Lariophagus* and *Heterospilis*

Lariophagus distinguendus is a small (1–3 mm) parasitic wasp (family Pteromalidae) which attacks larvae of the common granary weevil, *Sitophilus granarius*[22,29,30]. The weevils oviposit in grains of wheat, and a single larva develops inside the grain. The parasite stings the larva and deposits a single egg. An adult wasp emerges about 18 days later (at 25 °C). The wasp will parasitize a variety of host sizes[22]. The larva hollows out the grain as it develops. Size of weevil larvae is correlated with age, but the relationship is imprecise. For the present experiments we measured host size directly. The size of the weevil larva is strongly correlated with the diameter of its tunnel[37]. Tunnel diameter was measured on X-ray photographs[22,37].

Heterospilis prosopoidis is a small (1–3 mm) parasitic wasp in the family Braconidae. It has a similar life history to that of *Lariophagus* in that it lays a single egg on the larvae of weevils, attacking a variety of host sizes. It was studied here attacking the azuki bean weevil, *Callosobruchus chinensis*. At 30 °C, weevil development (egg to adult) takes about 21 days. Weevil larvae 11–19 days old are accepted by the wasp. Experiments with this system used weevil age as a measure of host size.

Van den Assem[22] had previously shown, using weevil age as a measure of size, that *Lariophagus* adjusts its sex ratio as a function of weevil size. Mostly males were produced in young weevils, mostly females in older weevils. His experiments were carried out by presenting wasps with weevils of only one age class. The experiments were also controlled to show that the observed sex ratio shift could not be due to differential, sex-specific mortality (even though young weevils produced a much higher death rate). The data also suggested that the wasp used a relative measure of host size.

We have repeated these earlier[22] experiments, using the more exact measure for host size. Figure 2*b* shows the results. As expected, sons are produced in small hosts, daughters in large. A second experiment consisted of presenting wasps with two host sizes, offered in equal abundance. The hosts were always 0.4 mm different in size and the wasp was presented with an alternating sequence (large, small, . . .). The data from these two experiments thus consist of a given host size in three situations: alone, smaller of two sizes and larger of two sizes. For example, we have 1.4-mm hosts alone, and paired with 1.0- or with 1.8-mm hosts. If we consider some host size *z*, it is clear that the three experiments (*z* small, *z* alone, *z* large) represent successively more left-shifted host size distribution curves. This corresponds to a shift from distribution ii to distribution i in Fig. 1*b*. If the wasp is adapted to alter its sex ratio on a short time scale, it should put relatively more daughters into host *z* as we move across the three treatments.

Curve *a* in Fig. 2 represents the sex ratio as a function of host size, when the host represented was the larger of the two hosts (for example, 1.0 was paired with 0.6). Note that every point lies below (is more female biased) the curve for the hosts when they were by themselves. Curve *c* (Fig. 2) shows the corresponding relationship for hosts where they were the smaller of the two hosts (for example, 1.0 was paired with 1.4). With a single exception, 0.7 mm, this curve lies above (more male biased) the other two curves. We have not provided statistical tests for the individual host sizes because the hypothesis relates to the general order of the curves. It is indeed remarkable that, even given small (10–30) to moderate (50–100) sample sizes, the curves preserve their theoretically predicted order, over the entire range of host sizes. This is the strongest evidence for the shift predicted by sex ratio theory.

Using host age as a measure of size, one of us (W.J.T.) carried out a similar set of experiments with *Heterospilis*. Presenting wasps with a single host size produced the relationship in Fig. 3*a*. Small hosts (11 days) were very male biased, whereas larger hosts were more female biased. A series of control experiments showed that this result could not be due to differential, sex-specific mortality. To see if these wasps could also adjust the sex ratio to a shifting host environment, we carried out a second set of experiments with just two host ages (11 and 15 days, or 15 and 19 days), presented in equal numbers. In comparison with the experiments in which the wasp received only one host size, the 11- and 15-day hosts now showed no altered sex ratio. Indeed, the 15-day hosts showed the same sex ratio when alone, or when combined with either 11- or 19-day hosts. Moderate to large sample size (88–1,024 emerging wasps per host size) makes these results quite reliable. Only one predicted sex ratio shift was seen. When 19-day hosts were presented alone, they gave a proportion of males of 0.36 ($n = 786$). However, when they were presented with equal numbers of 15-day hosts, the sex ratio dropped to 0.25 ($n = 88$). Such a shift is statistically significant (one-sided, Fisher exact test; $P = 0.03$) in the predicted direction.

We carried out a further set of experiments with *Heterospilis*. These involved 11- and 15-day hosts, presented to the wasp in various proportions (proportions of 11-day hosts 0.1, 0.3, 0.5, 0.7, 0.9). As can be seen in Fig. 3*b*, the wasp failed in 8 out of 10 cases to show any deviation (significant at the 0.05 level) from the sex ratio predicted from the experiments with only one host size. Clearly, there is no pattern of sex ratio shift with changing host proportions (at least for 11- and 15-day hosts). Although

Heterospilis does not immediately adjust its sex ratio to host size distributions, it is still possible that its offspring sex ratios are attuned to host distributions in a way predicted by the model. It may be that the host size distributions naturally encountered by *Heterospilis* do not vary greatly between wasp generations or that they vary within generations, and thus these wasps would not be expected to respond to short-term shifts in the distribution. At least for this species, a selection experiment or geographic comparison seems to be necessary to test the model further.

At least two other parasitic wasps which show a sex ratio shift with host size also show short-term adjustment of the sex ratio[23,24]. Both species overproduce daughters in large hosts, sons in small. Chewyreuv[23] studied parasitoid wasps of the genus *Pimpla* (family Ichneumonidae) which attack lepidopteran pupae. When he offered a female wasp a mixture of host pupae consisting of *Sphinx* (large) and *Piéris* (small), the wasp produced mostly females in the large, males in the small. However, when the *Piéris* pupae were alternated with those of *Vanessa* (still smaller), the same wasp produced female offspring in the former and males in the latter. Sandlan[24] also studied a pupal parasite of Lepidoptera, the wasp *Coccygomimus turionella*. Here, individual wasps confronted with a single host size gradually altered the sex ratio through time. If the single size was large, the sex ratio declined from being female biased to being more male biased. A small size showed the opposite change—male biased towards more female biased. All these sex ratio shifts are in the theoretical direction.

Fitness as a function of host size

As the sex ratio shift is towards daughters in larger hosts, a further test of the theory consists of determining whether a female gains more fitness by being large than does a male. Measuring individual fitness is very difficult, especially for males, where the value must reflect the relative ability to gain access to females. Using *Lariophagus*, a set of experiments was designed to investigate whether wasp size was correlated with the size of its host, and also whether increased wasp size became translated into increased individual fitness. Figure 4 shows the result of rearing almost 2,000 individuals. Clearly, larger weevils give rise to large wasps. Two facts about Fig. 4 are of particular interest. First, the relationship is different for males and females. From the same-sized host, males are on average smaller than the corresponding female. This presumably reflects the fact that males show a shorter developmental time to adulthood on any given host size[22]. Second, both curves show a steep linear increase over the weevil size range 0.7–1.3 mm. At 1.3 mm both show an abrupt transition. Average male size no longer increases with weevil size, whereas average female size goes up at a much reduced rate (slope of curve here is about 12% the previous slope).

Figure 5 shows that increased size translates into both increased male and increased female fitness. For females the measure is the lifetime production of eggs, for males it is length of life. This is not a complete measure of male fitness, because it gives no indication of male mating ability. Small males have no difficulty in getting females to accept them as mates[22]. Because females typically mate once for life, small males would seem to give adequate numbers of sperm, but we do not know how such small males fare in mate competition with larger males.

A similar set of experiments with female *Heterospilis* showed that lifetime fecundity increased about 20 times, comparing 11-day with 15-day hosts [1.3 offspring (s.d. = 2.87, $n = 61$) compared with 27.9 (s.d. = 8.38, $n = 46$)]. Interestingly, males from the same-aged hosts showed a much smaller fitness size effect, the measure here being females inseminated over a 3-h period (with 20 females available). Eleven-day males inseminated an average of 2.8 females (s.d. = 2.91, $n = 38$), compared with 8.3 (s.d. = 2.36, $n = 31$) for 15-day males. Although these data suggest that daughters gain more by being large than do sons, they are not conclusive. It would be very useful to know how individual size affects other fitness components such as female search ability or female ability to attack large hosts.

There is another type of data which bear on this fitness gain issue, at least for *Lariophagus*. If fitness within each sex is related to the size of the wasp, host sizes over the range 1.3–2.0 mm show very little change in relative female/male size (Fig. 4), whereas relative female/male size is increasing over the host size range 0.6–1.3 mm. We might then expect that host sizes above 1.3 would be treated quite similarly with respect to sex ratio. As shown in Fig. 2*b* (host size alone experiments), the proportion of sons drops rapidly with increasing host size until a host size of 1.3 mm. The sex ratio is about 0.17 males at this host size, but it does not become more female biased as we move to even larger hosts. The experiments presented in Fig. 2 show that the wasp can alter its sex ratio when a two-host environment consists of both hosts above 1.3 mm. However, these shifts are small relative to the alteration in sex ratio observed over a host size range of 0.6–1.3 mm. The sex ratio versus host age relationship measured earlier by van den Assem[22] shows a similar abrupt transition in sex ratio at the host age of 23–24 days. X-ray calibration shows that this age corresponds to a mean host size close to 1.3 mm.

Other hymenopteran systems

There are other parasitoid systems where a range of host sizes is attacked, and where sex ratio alters with host size[22–34]. The shift is always in the direction of more females from larger hosts. There are other hymenopteran systems which show similar trends. 'Trap nesting' (ref. 38) refers to free-living bees and wasps where the mother places food plus an egg in a cell constructed inside a crevice (for example, a hollowed-out twig). The size of the crevice places constraints on the cell size, the amount of food packed in and consequently the size of the resulting adult wasp[38,39]. Trap nesters have been much studied because they will commonly nest in soda-straws, holes drilled in wood blocks, or other man-made crevices[38,39]. Because of the importance of some trap-nesting bee species for agricultural pollination[39], many data now exist. Figure 6 shows the sex ratio (proportion of emerging males) as a function of hole diameter in artificial nesting material for 11 species of bees and wasps. With one exception, these data are from the book by Karl Krombein[38]. This documents no trap nester with a reverse sex ratio shift, although several species show evidence of no sex ratio shift for the hole sizes presented (several wasps of the genera *Stenodynerus* and *Ancistrocerus*). A reverse shift might be expected in species where the males show aggression and territoriality with respect to mating[35,40]. Here, increased size could

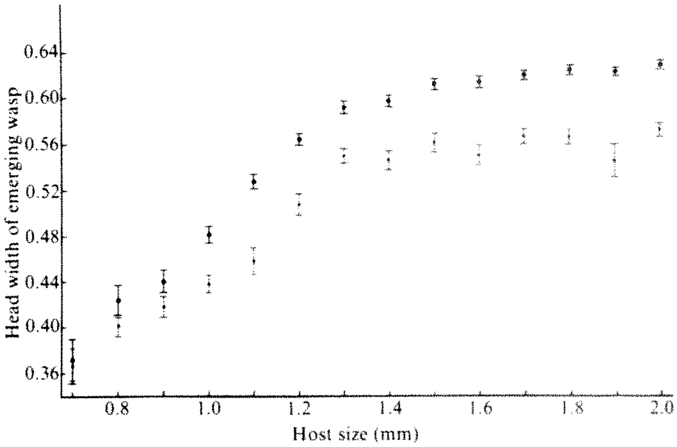


Fig. 4 Size of *Lariophagus* related to its host size (tunnel width). Means (±1 s.e.) of head width of males (lower curve) and females (upper curve) are plotted against host weevil size. Experimental details are given in Fig. 2 legend. Least squares regressions calculated on the means: (1) ♂, host sizes 0.7–1.3 mm, $y = 0.16 + 0.28x$, $r = 0.98$, $P < 0.001$. (2) ♂, host sizes 1.4–2.0 mm, correlation not significant at 0.05 level. (3) ♀, host sizes 0.7–1.3 mm, $y = 0.11 + 0.37x$, $r = 0.99$, $P < 0.001$. (4) ♀, host sizes 1.4–2.0 mm, $y = 0.54 + 0.044x$, $r = 0.92$, $P < 0.01$.

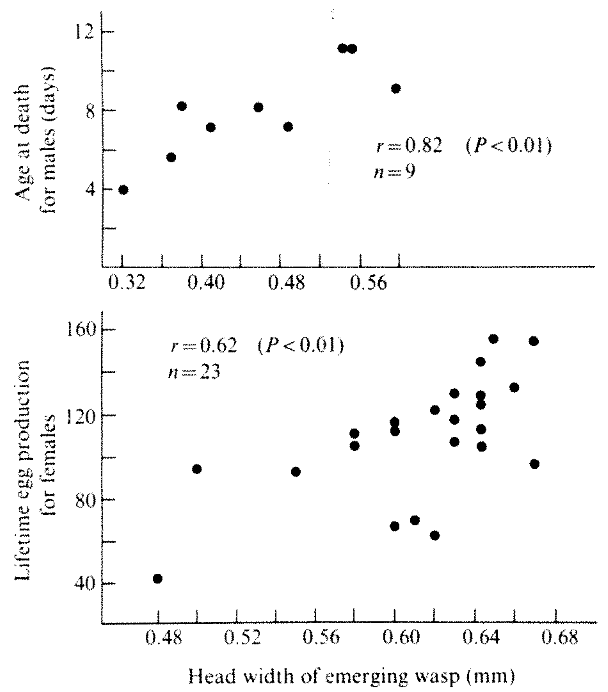


Fig. 5 Fitness as a function of individual size in *Lariophagus*. Both males and females gain by being large. Females were presented with an abundance of hosts. Experiments were run at 18 °C.

be translated into relatively greater male reproductive success[35]. It would be very interesting to know if any trap-nesting species shows geographical variation in the hole-size distribution. Such variaton should correlate with sex ratio shifts within a specified hole size, similar to those predicted (Fig. 1*b*, host size *z*) for parasitoids. However, the situation for trap nesters is somewhat more complicated than the simple model developed here. At least two other sex ratio effects occur with the hole-size shifts. There are often several cells placed in the same crevice—that is, a straw will contain a sequence of offspring. Because males typically emerge before females, it is common for the mother to put more daughters in the deep cells, more sons near the opening[38]. In this way, the sons can get out first without having to destroy the wasp in front. Except where different mothers may supersede eath other at a particular hole, the offspring in a given hole will be siblings. Some degree of cooperation may be expected[41], but the parents should benefit by using deep crevices for a mixture of sons and daughters. The other factor affecting sex ratio is that a son costs (in the sense of Fisher[18,21]) less than a daughter to produce. They require smaller cells (regardless of the diameter of the hole) and less food[38]. The equilibrium sex ratio should reflect this lowered cost through production of more males[20,21]. Even though these two factors complicate the situation, the hole-size effect, as illustrated in Fig. 6, still seems to be a dominating factor.

Discussion

Various biological systems in addition to Hymenoptera show sex ratio shifts of the type discussed here. These include pandalid shrimp[12,42], various orchids[43], entomogenous nematodes[44,45] and molluscs of the genus *Crepidula*[11]. In these cases, larger individuals typically become females, smaller ones male. Some of these species are sex reversers (sequential hermaphrodites), beginning life as a male and later changing into a female. Of course, many sex reversers go in the other direction; particularly among coral reef fish, species are known where individuals begin life as females, and become males when they are large enough to compete for a breeding territory[6-10,46].

As developed here and elsewhere[6-12,35,42,47,48], an understanding of the direction of the shift seems to depend on knowledge of which sex gains relatively more by being large. Whether the problem is the sex ratio to be produced by a wasp constrained to rear a small offspring or the 'choice' to develop into a small adult male or female imposed on a nematode in a crowded host[44,45], the principles underlying the force of natural selection remain the same.

Probably the largest gap in our knowledge is a clear demonstration of which sex gains relatively more by being large (except perhaps for some protogynous fish[6,7,46], where it seems fairly clear that males gain more). Because of the difficulty in measuring fitness with sufficient precision to test whether females gain more by being large, we propose here an alternative experiment. The data suggest that males also gain by being larger. If a laboratory culture can arrange for females not to gain fitness by being larger, the selection pressures will be to reverse the sex ratio shift. The culture conditions will then favour males in larger hosts, females in smaller.

E.L.C. thanks J. Bull and J. Werren for many years of discussion on problems of sex allocation. Ken Sandlan, J. Bull, Richard Green, J. van Ierse, John Endler and D. Davidson read and improved the paper. E.L.C. and W.T.J. were supported by grants from the US NSF. J. v.d. A. and R.L.L-d.H. were supported by the Netherlands Organization for the Advancement of Pure Science (ZWO), grant 89-177.

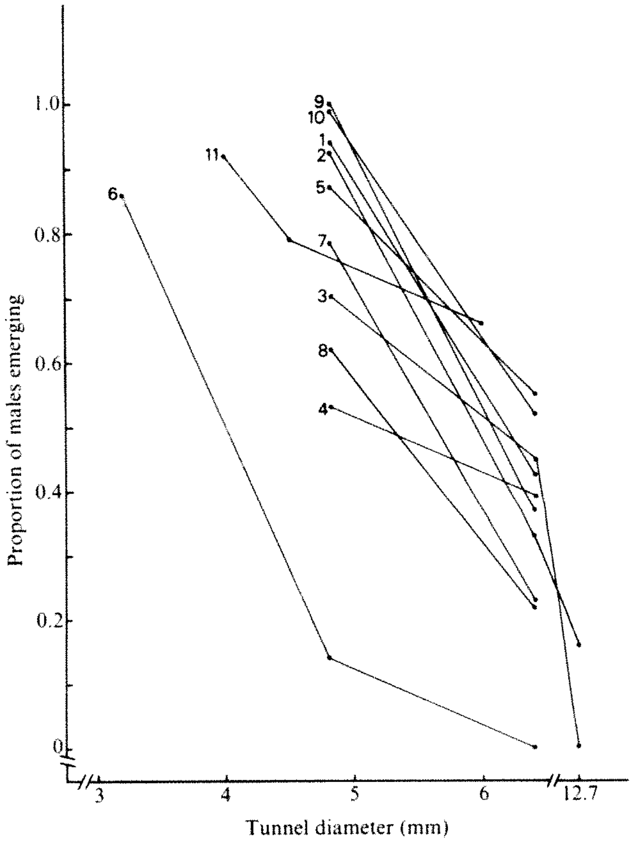


Fig. 6 Sex ratio at emergence related to tunnel diameter of nesting material for trap-nesting Hymenoptera. Data points refer to diameter of nesting material used in sampling. Lines connecting points are simply to keep species distinct. All sex ratio shifts are statistically significant at the 0.05 level (one-tailed, Fisher exact test) or better. Species 1–10 are from ref. 38, species 11 from ref. 39. Species as follows: 1, *Euodynerus foraminatus*; 2, *Euodynerus megaera*; 3, *Euodynerus schwarzi*; 4, *Pachodynerus erynnis*; 5, *Ancistrocerus antilope*; 6, *Ancistrocerus tigris*; 7, *Symmorphus canadensis*; 8, *Tryarpilum tridentatum* (North Carolina); 9, *Tryarpilum johannis*; 10, *Osmia lignaria*; 11, *Megachile rotundata*. Species 1–9 are wasps, 10 and 11 are bees.

Received 4 August; accepted 22 September 1980.

1. Charnov, E. L. *Am. Nat.* **113,** 465–480 (1979).
2. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
3. Williams, G. C. *Sex and Evolution* (Princeton University Press, 1975).
4. Charnov, E. L., Maynard Smith, J. & Bull, J. J. *Nature* **263,** 125–126 (1976).
5. Charnov, E. L. *Proc natn. Acad. Sci. U.S.A.* **76,** 2480–2484 (1979).
6. Warner, R. R. *Am. Nat.* **109,** 61–82 (1975).
7. Warner, R. R., Robertson, D. R. & Leigh, E. G. *Science* **190,** 633–638 (1975).
8. Leigh, E. G. Jr, Charnov, E. L. & Warner, R. R. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3656–3660 (1976).
9. Ghiselin, M. T. *Q. Rev. Biol.* **44,** 189–208 (1969).
10. Ghiselin, M. T. *The Economy of Nature and the Evolution of Sex* (University of California Press, 1974).
11. Hoagland, K. E. *Malacologia* **17,** 365–391 (1978).
12. Charnov, E. L. *Am. Nat.* **113,** 715–734 (1979).
13. Lloyd, D. G. *N. Z. J. Bot.* **17,** 595–606 (1979).
14. Charlesworth, B. & Charlesworth, D. *Am. Nat.* **112,** 975–997 (1978).
15. Willson, M. F. *Am. Nat.* **113,** 777–790 (1979).
16. Bawa, K. S. *Ann. Rev. Ecol. Syst.* **11** (in the press).
17. Maynard Smith, J. *Am. Scient.* **64,** 41–45 (1976).
18. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
19. Hamilton, W. D. *Science* **156,** 477–488 (1967).
20. Trivers, R. L. & Hare, H. *Science* **191,** 249–263 (1976).
21. Charnov, E. L. *Am. Nat.* **112,** 317–326 (1978).
22. Assem, J. van den *Neth. J. Zool.* **21,** 373–402 (1971).
23. Chweyreuv, I. *C. r. Séanc. Soc. Biol.* **74,** 695–699 (1913).
24. Sandlan, K. *Ecol. Ent.* **4,** 365–378 (1979).
25. Clausen, C. D. *J. N. Y. ent. Soc.* **47,** 1–9 (1939).
26. Holdaway, F. T. & Smith, H. F. *Aust. J. exp. Biol. med. Sci.* **10,** 247–159 (1933).
27. Brunson, M. H. *Science* **86,** 197 (1937).
28. Nozato, K. *Kontyu* **37,** 134–146 (1969).
29. Assem, J. van den *Neth. J. Zool.* **20,** 329–352 (1970).
30. Assem, J. van den in *Sluipwespen in relatie tot hun gastheren* (eds Klomp, H. & Wiebes, J. T.) 64–96 (Centrum voor Landbouwpublicaties en documentatie, Wageningen, 1979).
31. Arthur, A. P. & Wylie, H. G. *Entomophaga* **4,** 297–309 (1959).
32. Ryan, R. B. & Rudinsky, J. A. *Can. Ent.* **94,** 748–763 (1962).
33. Berry, P. A. *J. Econ. Ent.* **32,** 717–721 (1939).
34. Kishi, Y. *Appl. ent. Zool.* **5,** 126–132 (1970).
35. Trivers, R. L. & Willard, D. E. *Science* **179,** 90–92 (1973).
36. Hamilton, W. D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.), (Academic, New York, 1979).
37. Kirkpatrick, R. L. & Wilbur, D. A. *J. econ. Ent.* **58,** 979–985 (1965).
38. Krombein, K. V. *Trap Nesting Wasps and Bees* (Smithsonian, Washington, DC, 1968).
39. Stephen, W. P. & Osgood, C. E. *J. econ. Ent.* **58,** 965–968 (1965).
40. Baroni Urbani, C. in *Social Insects* Vol. 1 (ed. Hermann, H. R.) 19–121 (Academic, New York, 1979).
41. Hamilton, W. D. *J. theor. Biol.* **7,** 1–52 (1964).
42. Charnov, E. L., Gotshall, D. & Robinson, J. *Science* **200,** 204–206 (1978).
43. Gregg, K. *Selbyana* **1,** 101–113 (1975).
44. Peterson, J. J. *J. Nematol.* **4,** 83–87 (1972).
45. Poinar, G. O. Jr *Nematodes for Biological Control of Insects* (CRC Press, Boca Raton, Florida, 1979).
46. Shapiro, D. Y. *Adv. Study Behav.* **10,** 43–102 (1979).
47. Charnov, E. L. & Bull, J. J. *Nature* **266,** 828–830 (1977).
48. Freeman, D. C., Harper, K. T. & Charnov, E. L. *Oecologia* (in the press).

Isolation of a gene from *Drosophila* by complementation in yeast

Steven Henikoff[*], Kelly Tatchell[†], Benjamin D. Hall[†] & Kim A. Nasmyth[†]

[*] Department of Zoology and [†] Department of Genetics, University of Washington, Seattle, Washington 98195

Transformation of mutant yeast cells by cloned genomic DNA from a higher eukaryote has made it possible to isolate a Drosophila *DNA sequence that complements a yeast adenine-8 mutation. A 0.8-kilobase poly(A)-containing RNA is transcribed from the cloned* Drosophila *segment in transformed yeast cells and can account for functional expression of the gene.*

PHYSICAL analysis of a large number of genes from animals has yielded an overall picture of the eukaryotic gene as a complex entity, sandwiched between long untranslated regions and interrupted several times by sequences which are not present in the mature mRNA. The procedures for isolating animal genes generally depend on specific nucleic acid sequence probes obtained by virtue of an enriched content of a particular mRNA in some specialized tissue or developmental stage. Consequently, the animal genes that have been most studied are seldom expressed in more than a few cell types. It is possible that some of the structural features attributed to higher eukaryotic genes are only typical of genes for differentiated functions. It is not known whether genes for cellular metabolism and cell division show these same features because of the absence of methods for isolating them from animal cells.

We now describe a gene substitution method for the isolation and molecular genetic analysis of metabolic genes from higher eukaryotes. The method is based on functional complementation of mutant yeast cells by recombinant plasmids from a random genomic library of *Drosophila* DNA fragments inserted into a yeast vector. By this procedure we have cloned a *Drosophila* gene that complements a mutation at the yeast adenine-8 locus.

Yeast transformation with pYF DNA

The cloning of yeast genes by complementation of yeast mutations has been successfully applied for genes affecting both cell metabolism[1-3] and the cell cycle[4]. In our experiments adenine auxotrophs were chosen because all organisms are thought to synthesize purines by a universal *de novo* pathway. We constructed recombinant plasmids by inserting *Drosophila* genomic fragments into a vector that could be replicated and selected for in both *Escherichia coli* and yeast. *Drosophila melanogaster* embryonic DNA was partially digested with *Sau*3A, a restriction endonuclease that recognizes GATC, cutting 5′ to the G to leave single-stranded GATC ends[4]. A mixture of such *Sau*3A fragments with a mean size of ~5 kilobases was then ligated into the single *Bam*HI restriction site (G↓GATCC) of the plasmid YEp13 (ref. 2), and the ligation mix was used to transform *E. coli* to give a quasi-random pool of cloned *Drosophila* fragments. Figure 1 shows the structure of a pYF plasmid. The bacterial ampicillin resistance gene of pBR322 allows selection in *E. coli* and the yeast *LEU*2 gene allows selection in *Saccharomyces cerevisiae*. The presence of yeast 2-μm circle DNA permits plasmid replication in yeast without integration.

A plasmid DNA pool containing many YEp13-*Drosophila* recombinants was purified from *E. coli* and used to transform mutant yeast strains[4,7]. To determine the adequacy of selective conditions, DNA from an identically constructed yeast DNA pYe pool was used to transform aliquots of mutant cells. Cells mutant for both *leu*2 and the test marker were plated in the absence of both leucine and adenine, so that only cells carrying a plasmid with both *LEU*2 and the Ade$^+$ gene being tested should have grown. We tested six purine auxotrophic strains, *ade*1–*ade*4, *ade*5,7 and *ade*8. Ade$^+$ prototrophs were found in every case for cells transformed with the pYe pool, except for *ade*2. (The pYe pool was constructed with DNA from an *ade*2 mutant strain.) For cells transformed with the pYF pool, Ade$^+$ prototrophs were found in one case out of six—*ade*8.

For *ade*8, six independent prototrophs were found from the pYF pool—these were picked up at a frequency of about 1 per 10^4 Leu$^+$ transformants whereas prototrophs from the pYe pool were picked up about an order of magnitude more frequently. Because the genome size of *Drosophila* is about 10 times that of yeast, one would expect the fly *ADE8* gene to be present ~1/10th as often in the pYF pool as its counterpart in the pYe

0028–0836/81/020033—05$01.00

pool. These pYF and pYe prototrophs were shown to be genuine *ADE*8 transformants rather than revertants or contaminants by the behaviour of these cells when grown without selection. YEp13 plasmids are lost at a frequency of ~1% per cell generation from yeast so that after several generations of growth on complete medium, a large fraction of cells will be leucine-requiring. Such cells should also be *ade*8 if *ADE*8 is carried on the recombinant plasmid. By this test, all *ADE*8 transformants, whether derived from the pYe or from the pYF pool, were found to be carried on a *LEU*2 plasmid—that is, loss of Leu$^+$ was always accompanied by adenine auxotrophy.

Restriction analysis of *ADE*8 plasmids

DNA was isolated from the six pYF and from six pYe *ADE*8 transformants grown up with adenine selection. DNA samples were used to transform *E. coli*, from which single ampicillin-resistant clones were grown and plasmid DNA extracted. Restriction endonuclease analysis of these plasmids is summarized in Fig. 2. The six pYF plasmids showed three different but partially overlapping restriction patterns, and the six pYe plasmids showed two. Apparently the smallest pYF insert is contained entirely within the other two: it has the same orientation as the second longest. This was confirmed by hybridization analysis using one of the plasmids as a probe (data not shown). No cross homology was detected, however, between the pYF and the pYe *ADE*8 plasmid inserts.

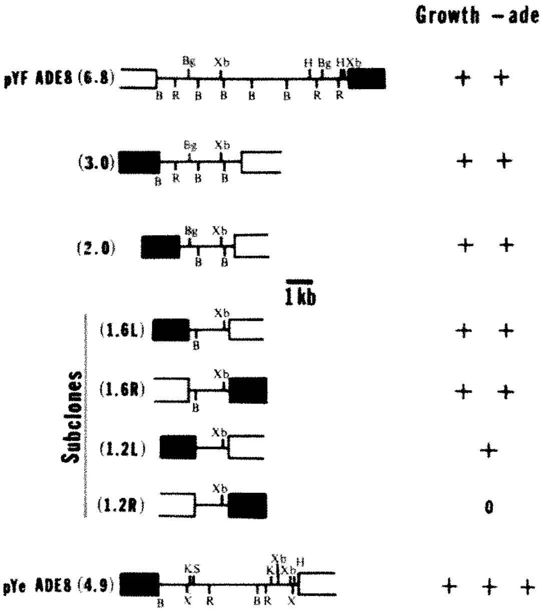


Fig. 2 Restriction endonuclease analysis of *ADE*8 plasmid inserts and relative growth with adenine selection of *ade*8 cells carrying *ADE*8 plasmids. Orientation of the inserts within the YEp13 vector is indicated by the solid panel, which corresponds to the region on the left hand side of the *Drosophila* segment in Fig. 1. Subclones were constructed by ligating the gel-purified 1.6-kilobase *Bgl*II–*Bam*HI and the 1.2-kilobase *Bam*HI–*Bam*HI fragments, derived from pYF *ADE*8(3.0) and pYF *ADE*8(2.0) respectively, into the *Bam*HI site of YEp13 (ref. 2). All pYF *ADE*8 inserts appeared to overlap with one another but showed no similarity to the pYe *ADE*8 inserts, one example of which is shown. The following enzymes were used: *Bam*HI (B), *Bgl*II (Bg), *Hin*dIII (H), *Kpn*I (K), *Eco*RI (R), *Sal*I (S), *Xba*I (Xb), and *Xho*I (X). Growth rates were determined by colony size estimation for initial isolates and subclones tested in the same yeast strain used for the initial selection (5028-1Dα ade8-18 leu2-1 met2 his4-4 lysX gal cup^s). Transformants were tested on plates minus leucine and on plates minus adenine. No differences in colony size were noted for any transformant using leucine selection. Growth using adenine selection was estimated from colony size after 3 days at 29 °C. kb, Kilobase.

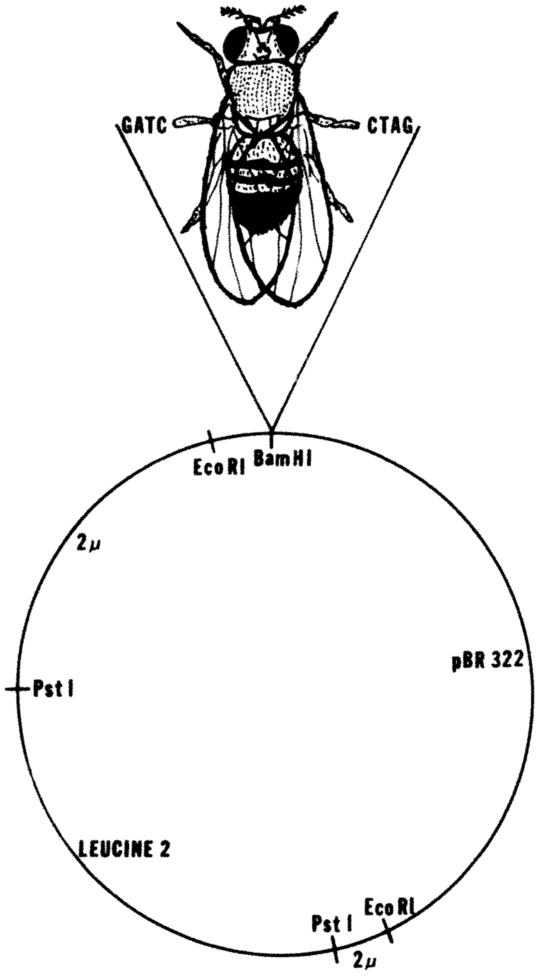


Fig. 1 Schematic map of pYF plasmid DNA derived from the bacterial plasmid pBR322, yeast 2-μm circle DNA, the yeast *LEU*2 gene and *Drosophila* embryonic DNA. Fresh *D. melanogaster* (wild-type 'Sevelin') embryos were thoroughly washed and dechorionated 8–15 h after egg deposition, and 2,000*g* nuclei were prepared[5]. Microbial contamination was checked by plating out and was found to be insignificant. Nuclei were lysed in 10 mM Tris *p*H 7.9, 4 mM EDTA, 1% Sarkosyl, and DNA was banded in CsCl and dialysed. Partial digestion of this DNA with *Sau*3A, ligation into YEp13 (ref. 2) and transformation was performed as described elsewhere[4]. A pool of *E. coli* transformants (3×10^5) with mean *Drosophila* insert size of 4–5 kilobases was obtained and was enriched by selection for ampicillin resistance and tetracycline sensitivity[6]. A log-phase culture inhibited with chloramphenicol was used for plasmid DNA isolation and purification[4].

To demonstrate directly the presence of an *ADE*8 gene, *ade*8 *leu*2 yeast cells were transformed with each one of the plasmids being analysed. In every case, both *ade*8 and *leu*2 were simultaneously complemented, proving that an *ADE*8 gene is present within each plasmid insert (data not shown). Using leucine selection, all such transformants grow about as well as wild-type cells regardless of which plasmid is used. However, when adenine is absent, an interesting difference appears. Whereas yeast cells carrying pYe-derived plasmids grow about as rapidly as do wild type, those with any of the pYF plasmids grow at about two-thirds the normal rate. This reduced rate of growth attributable to the pYF *ADE*8 gene suggests that the *Drosophila* *ADE*8 gene functions less effectively in yeast than its yeast functional counterpart.

Hybridization of pYF *ADE*8 to *Drosophila* DNA

An unlikely artefact that could account for complementation of a yeast auxotroph by pYF plasmids is microbial contamination of the *Drosophila* DNA, although precautions were taken to prepare very clean dechorionated embryos. Southern transfer experiments can rule out this possibility by demonstrating that the pYF plasmid inserts are derived from *Drosophila*[8]. Figure 3 shows an agarose gel on which *Drosophila* genomic DNA was electrophoresed after cutting with *Bam*HI restriction enzyme. In another lane, an amount of pYF *ADE*8(2.0) representing a genome equivalent was also digested with *Bam*HI and electrophoresed. This gel was blotted and probed with the 1,200-base pair fragment produced by *Bam*HI digestion of pYF *ADE*8(2.0) and labelled with ^{32}P by nick-translation. Figure 3*c*, *d* shows that a 1,200-base pair band is present in both lanes which indicates that the pYF *ADE*8 inserts are derived from *Drosophila*. Also, the extent of hybridization is the same in both lanes which suggests that the sequence is single-copy.

This result has been confirmed using *Drosophila* genomic DNA from embryos of different strains and from polytene fat body cut with *Eco*RI, *Xba*I and *Bgl*II. In each case, a single band of the expected size was hybridized to the extent expected for a sequence present once per haploid genome (data not shown). Thus we conclude that the pYF sequence complementing *ade*8 is derived from *Drosophila* and is single-copy.

A *Drosophila* sequence present once per haploid genome in polytene tissue can be localized unambiguously on salivary chromosomes. Figure 4 shows the result of an *in situ* hybridization experiment using ^{3}H cRNA to pYF *ADE*8(6.8) as a probe. A single site, 27C on chromosome arm 2L, is labelled. Even exposures 10 times longer than for the example in Fig. 4 do not show labelling of any other site. The region around the labelled site is not one of extensive genetic characterization[13]. However, in screening for adenosine auxotrophs on chromosome 2, Naguib and Nash (personal communication) have isolated two non-allelic *Drosophila* mutations, and both map to this region. It is tempting to speculate that one of these auxotrophic mutations blocks the reaction catalysed by the *Drosophila* enzyme that is coded for by pYF *ADE*8. In yeast, adenine-8 is presumed to be the gene for the enzyme glycinamide ribotide (GAR) transformylase[14]. It will be interesting to determine whether the same enzyme activity is specified by this sequence in *Drosophila*.

Transcription of pYF *ADE*8

Because the *Drosophila* segments that complement *ade*8 were carried on plasmids that also included bacterial and yeast DNA, we wondered whether or not *ADE*8 gene expression depends entirely on *Drosophila* DNA. For example, transcription and translation could conceivably initiate or terminate within flanking vector DNA so that the *Drosophila* sequence would supply only structural information. We found that this was not likely.

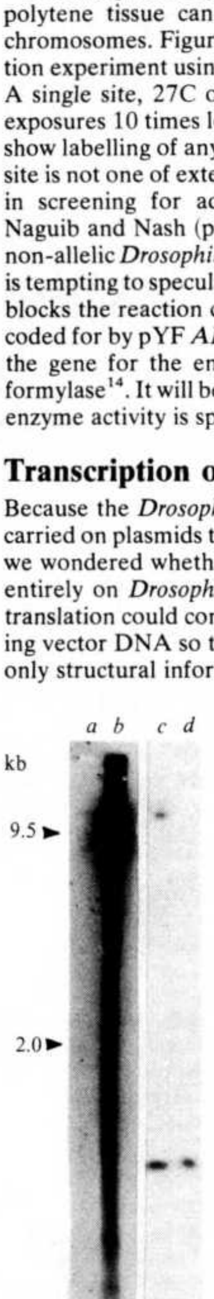


Fig. 3 Hybridization of *Drosophila* genomic DNA with pYF *ADE*8 probe. *a*, An aliquot (3 ng) of pYF *ADE*8(2.0) DNA was digested with *Bam*HI and electrophoresed on a 0.9% agarose gel[9] adjacent to *b*: 10 μg *Bam*HI-digested *Drosophila* embryonic DNA. After staining with ethidium bromide and photographing, the DNA was transferred to nitrocellulose and hybridized[10] with the 1,200-base pair *Bam*HI fragment partially purified (95%) from pYF *ADE*8(2.0) and labelled with ^{32}P by nick-translation. The nick-translation reaction mixture (50 μl) contained 10 μCi [α-^{32}P]dCTP (NEN); 4 μM dATP, dGTP and dTTP; 50 mM Tris-HCl *p*H 8; 5 mM $MgCl_2$; 50 μg per ml bovine serum albumin; ~1 μg DNA; 3 units DNA polymerase I (Boehringer) and 75 pg DNase I (Worthington). *c*, *d*, Autoradiograms of tracks *a* and *b*, respectively.

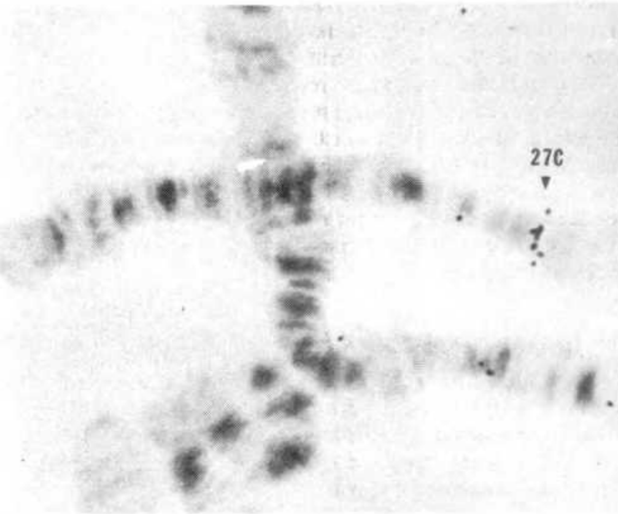


Fig. 4 Hybridization of pYF *ADE*8(6.8) to polytene chromosomes *in situ*. ^{3}H-labelled copy RNA[11] was hybridized as described elsewhere[12] to *gt* w^a/gt^{X11} prepupae. Exposure was for 3 days.

The gene functioned at the same reduced rate for all three different plasmid types, even though their inserts had different orientations. More direct evidence that *ADE*8 gene expression depended on *Drosophila* DNA and not flanking vector sequences came from an investigation of transcription on a pYF *ADE*8 plasmid.

If pYF *ADE*8 complements the yeast *ade*8 mutation by making a protein, then one would expect to find poly(A)-containing RNA hybridizable to *Drosophila ADE*8 DNA in cells transformed with pYF *ADE*8 plasmids. Adenine-8 mutant cells that carry pYF *ADE*8(2.0) were grown with adenine selection and poly(A)$^+$ RNA was isolated. An aliquot of this RNA was denatured in dimethyl sulphoxide (DMSO) and glyoxylated, then electrophoresed on agarose and transferred to nitrocellulose as described elsewhere[15]. This 'Northern' transfer was probed with the 1.5-kilobase fragment derived from the pYF *ADE*8(2.0) insert by digestion with *Bgl*II and *Xba*I and labelled with ^{32}P by nick-translation. Figure 5 shows the resulting autoradiogram. Two prominent bands are seen, one at 0.8 kilobases and another at 2.2 kilobases, with a few minor bands and some heterogeneous labelling also visible. The 2.2-kilobase transcript is very likely to have been read from vector sequences in part, as it seems to be longer than the entire 2-kilobase *Drosophila* insert in the plasmid that was carried by the cells from which the RNA was derived. Experiments described below demonstrate that the 0.8-kilobase transcript is derived entirely from *Drosophila* DNA, and that it is responsible for *ADE*8 function.

One way to determine which transcript provides *ADE*8 function is physically to map the transcript on the plasmid, then compare this map with a functional map determined by testing partially deleted subclones for *in vivo* function. We used S_1 nuclease protection of transcripts by various fragments to map the transcribed region[18]. Poly(A)$^+$ RNA used in the above Northern blot experiment was hybridized to the 1,500-base pair *Bgl*II–*Xba*I fragment in conditions that allow RNA–DNA but not DNA–DNA duplex formation. In parallel reaction mixtures, the DNA fragment was also first digested with either *Bam*HI or *Ava*II to give fragments that were 350 or 375 base pairs shorter, as shown in Fig. 6*a*. These mixtures were then digested with S_1 nuclease, electrophoresed on an agarose gel and transferred to nitrocellulose. This blot was probed with pYF *ADE*8(2.0) labelled with ^{32}P by nick-translation and autoradiographed—the result is shown in Fig. 6*a*. All three fragments are protected by transcript to give a major band at ~0.8 kilobases and a minor band at ~0.7 kilobases. Both these bands are the same size when either the *Bgl*II–*Xba*I fragment or

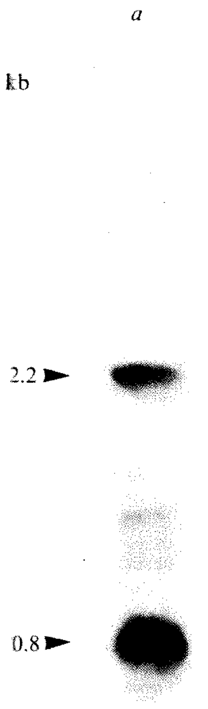


Fig. 5 Hybridization of pYF *ADE*8 DNA to a Northern blot of poly(A)-containing RNA. Log-phase *ade*8 *leu*2 yeast cells carrying *a*, pYF *ADE*8(2.0) and *b*, pYE *ADE*8(4.9) were grown with adenine selection. RNA was extracted[16] and the poly(A)$^+$ fraction isolated[17]. Samples (10 μg) were glyoxylated in 80% DMSO at 65 °C, electrophoresed, transferred to nitrocellulose and hybridized using the procedure of Thomas[15]. The probe used was the purified *Bgl*II–*Xba*I fragment derived from the insert of pYF *ADE*8(3.0) purified on a 1.1% agarose gel[9], electroeluted and labelled with ^{32}P by nick-translation.

the *Bam*HI–*Xba*I fragment is used for hybridization but are reduced in size by about 25 nucleotides when the *Ava*II–*Xba*I fragment is used. This result allows assignment of both transcripts to a region stretching from the *Bam*HI site to a point 0.8 kilobases to the right for the major transcript and 0.7 kilobases for the minor. This minor transcript is probably produced as a result of the S_1 digestion procedure because in milder conditions it is absent (data not shown). An S_1 nuclease-protected region of 0.8 kilobases corresponds well to the 0.8-kilobase transcript detected on the Northern blot. Failure to detect the 2.2-kilobase transcript here suggests that it is complementary to only a small part at one end of the *Bgl*II–*Xba*I fragment, and due to its small size, escapes detection. This experiment also suggests that there are no intervening sequences because the protected fragments are all of the same size in this neutral gel, as they were in an alkaline denaturing gel run in parallel (not shown).

The orientation of the 0.8-kilobase transcript was determined in a similar S_1 nuclease protection experiment using fragments labelled at one end only. Two different fragments were used (Fig. 6*b*). Starting with the 1.2-kilobase fragment liberated by *Bam*HI digestion, one labelled fragment (A) was made by cutting with *Ava*II, labelling the 3′ ends and then cutting with *Xba*I. The other labelled fragment (B) was made similarly but without the *Ava*II digestion. Ignoring the small fragments produced, fragment A is labelled at the *Ava*II site on its 3′ end and fragment B is labelled at the leftmost *Bam*HI site on its 3′ end. Only if the transcript is oriented with its 5′ end on the left and covers the labelled 3′ end restriction site will that site be protected from S_1 nuclease digestion. This is the case for the *Ava*II site as seen in the autoradiogram of a polyacrylamide gel (Fig. 6*b*). For fragment B, S_1 nuclease protection is also found, although not as much as for fragment A. Partial protection of the *Bam*HI site shows that the 5′ end of the 0.8-kilobase transcript is within a few nucleotides of this site.

Behaviour of subcloned pYF *ADE*8 fragments *in vivo*

Our evidence that the 0.8-kilobase transcript made in yeast starts and stops within the *Drosophila* sequences of pYF *ADE*8(2.0) does not unequivocally demonstrate that this

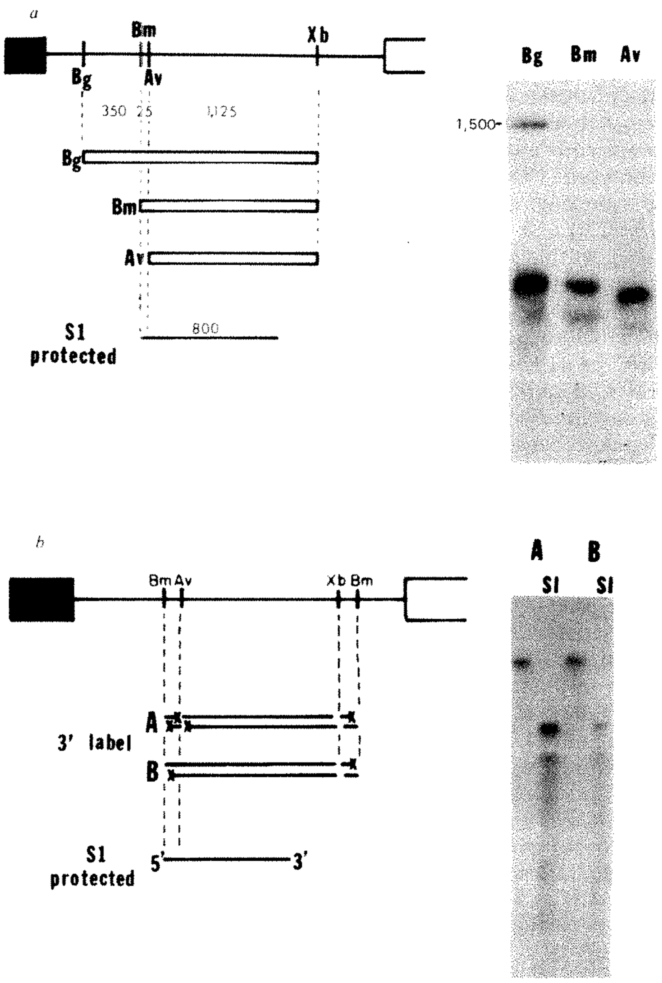


Fig. 6 Mapping of transcription by S_1 nuclease protection of restriction fragments derived from pYF *ADE*8(2.0). Purified restriction fragments were hybridized to poly(A)-containing RNA from cells carrying pYF *ADE*8(2.0) and S_1 nuclease-digested according to Holmgren *et al.*[19]. Each hybridization mixture (20 μl) contained 80% formamide, 0.3 M NaCl, 1 mM EDTA, 10 mM PIPES *p*H 6.8, 10–40 μg RNA and 80–160 ng DNA. Samples were kept at 70 °C for 10 min then transferred to 53 °C for overnight incubation. These were mixed with 20-μl aliquots of 0.3 M NaCl, 6 mM $ZnCl_2$, 100 mM Na-acetate, *p*H 4.0, and 50,000 units per ml S_1 nuclease, incubated for 5 min at 37 °C, extracted with phenol and ethanol-precipitated. *a*, Protection of unlabelled fragments. Hybrids were electrophoresed on a neutral 1.5% agarose gel, blotted onto nitrocellulose, hybridized with pYF *ADE*8(2.0) which was labelled with ^{32}P by nick-translation, and autoradiographed. The band at 1,500 base pairs is apparently due to DNA–DNA reassociation of the *Bgl*II–*Xba*I fragment. *b*, Protection of 3′ end-labelled DNA fragments. Before hybridization, the 1.2-kilobase *Bam*HI–*Bam*HI fragment purified from pYF *ADE*8(2.0) was either digested with *Ava*II (A) or used uncut (B) to fill the 3′ ends (indicated by x) with ^{32}P-labelled deoxynucleotide triphosphates. Both samples were digested with *Xba*I and aliquots used for hybridization and S_1 nuclease digestion. Hybrids were then redissolved in 0.1 M NaOH and electrophoresed on 4% acrylamide gels containing 7 M urea[20]. The resulting autoradiogram is shown at the lower right. For each labelled fragment A and B, an aliquot of the fragment itself was run to the left of the S_1 nuclease-digested sample (S1). Because the specific activities of the A and B large fragments differed, their concentrations were adjusted for the gel to give the same amount of radioactivity in both untreated large fragments. For the S_1 nuclease-treated samples, the same relative adjustment in concentration was made so that these gel tracks could be quantitatively compared for A and B. The 0.7-kilobase minor band seen after S_1 nuclease digestion is probably a digestion artefact (see text). Restriction sites are *Bgl*II (Bg), *Bam*HI (Bm), *Ava*II (Av) and *Xba*I (Xb). The orientation of the poly(A)$^+$ transcript is shown.

sequence is responsible for *ade*8 complementation. For example, it is possible that a less abundant transcript accounts for *ADE*8 function. However, the following *in vivo* tests of subcloned fragments strongly suggest that the 0.8-kilobase transcript mapped above is the functional message. As shown in Fig. 2, we tested fragments subcloned into the *Bam*HI site of YEp13 in both orientations. The 1.6-kilobase fragment made by *Bgl*II digestion and *Bam*HI partial digestion of pYF *ADE*8(2.0) showed *ADE*8 complementation activity equivalent to that of the original plasmids when subcloned in either orientation. The 1.2-kilobase fragment made by complete *Bam*HI digestion of pYF *ADE*8(2.0) showed reduced *ADE*8 activity in one orientation and no activity at all when subcloned in the other orientation. Because the 350 base pairs between the *Bgl*II site and the leftmost *Bam*HI site are the only difference between the subclones in either orientation, the structural gene apparently lies between the two *Bam*HI sites with a control region just to the left. This mapping of *Drosophila* adenine-8 based on function agrees well with the transcriptional mapping, because the entire transcript, except for possibly a few nucleotides at the 5′ end, is read from a region that can function in a subclone. It seems likely that substitution of sequences just upstream of the structural gene, perhaps even replacing the initiation site, has reduced or eliminated gene expression. Therefore, we conclude that the 0.8-kilobase poly(A)$^+$ transcript is the message in yeast and that the *Drosophila* segment from which it is transcribed is entirely responsible for complementation of *ade*8.

That this transcript functions as message is also apparent from a preliminary analysis of its nucleotide sequence. An open reading frame has been found to cover most of the transcribed region (C. E. Furlong and S.H., unpublished results). Presumably, this sequence codes for the enzyme GAR transformylase from *D. melanogaster.*

Concluding remarks

We have shown that yeast transformation can be used to isolate a gene from a higher eukaryote and to determine that transcription of the gene corresponds to its functional expression. Of the six adenine pathway genes investigated, only *Drosophila* adenine-8$^+$ was isolated in yeast by our procedure. Several explanations can be offered for the failure to find complementation for mutations in the other five yeast genes, including cases in which multi-enzyme genes in yeast may not be linked in *Drosophila* (for example, *ade*3 and *ade*5,7), transcripts may not be correctly processed or efficiently translated in yeast, or enzyme products may not be sufficiently functional or stable. Our characterization of *Drosophila* adenine-8$^+$ shows that it is transcribed in yeast to make a product that can substitute for the yeast *ADE*8 gene product. This suggests that the requirements for functional expression of this gene are very much the same in insects and fungi. As this is one of the first higher eukaryotic genes of a general metabolic nature to be analysed, it is possible that such genes will generally be similar to those in yeast. The gene substitution method used may prove useful for further inter-kingdom comparisons.

We thank Charles Laird for support and encouragement, Carol Sandstrom for help with embryo collection, Jim Sloan for testing yeast strains, Pat Thomas for providing the RNA blotting protocol and Seymour Fogel for the *ade*8 strain. This work was supported by grants from NIH to B. D. H. and to C. Laird, by a Damon Runyon-Walter Wichell Cancer Fund postdoctoral fellowship to K. T. and by a Jane Coffin Childs Memorial Fund for Medical Research postdoctoral fellowship to K. A. N.

Received 21 August; accepted 29 October 1980.

1. Hinnen, A., Farabaugh, P., Ilgen, C., Fink, G. R. & Friesen, J. *ICN-UCLA Symp. molec. cell. Biol.* **14,** 43–50 (1979).
2. Broach, J. R., Strathern, J. N. & Hicks, J. B. *Gene* **8,** 121–133 (1979).
3. Williamson, V. M., Bennetsen, J., Young, E. T., Nasmyth, K. & Hall, B. D. *Nature* **283,** 214–216 (1980).
4. Nasmyth, K. A. & Reed, S. I. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2119–2123 (1980).
5. Elgin, S. C. R. & Miller, D. W. in *The Genetics and Biology of Drosophila* Vol. 2a (eds Ashburner, M. E. & Wright, T. R. F.) 145–150 (Academic, New York, 1978).
6. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
7. Beggs, J. D. *Nature* **275,** 104–108 (1978).
8. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
9. Helling, R. B., Goodman, H. M. & Boyer, H. W. *J. Virol.* **14,** 1235–1244 (1974).
10. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3683–3687 (1979).
11. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. *Cell* **3,** 315–325 (1974).
12. Henikoff, S. & Meselson, M. *Cell* **12,** 441–451 (1977).
13. Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* Publ. No. 627 (Carnegie Institution of Washington, 1968).
14. Woods, R. & Jackson, I. *Biochem. biophys. Res. Commun.* **53,** 787–793 (1973).
15. Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5201–5205 (1980).
16. Schultz, L. D. *Biochemistry* **17,** 750–758 (1978).
17. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1408–1412 (1972).
18. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732.
19. Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. *Cell* **18,** 1359–1370 (1979).
20. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).

Intervening sequences in a plant gene—comparison of the partial sequence of cDNA and genomic DNA of French bean phaseolin

S. M. Sun[†], **J. L. Slightom**[*†] **& T. C. Hall**[‖]

Departments of Horticulture and Genetics*, University of Wisconsin, Madison, Wisconsin 53706

A plant gene coding for the major storage protein (phaseolin, G1-globulin) of the French bean was isolated from a genomic library constructed in the phage vector Charon 24A. Comparison of the nucleotide sequence of part of the gene with that of the cloned messenger RNA (cDNA) revealed the presence of three intervening sequences, all beginning with GT and ending with AG. The 5′ and 3′ boundaries of intervening sequences IVS-A (88 base pairs) and IVS-B (124 base pairs) are similar to those described for animal and viral genes, but the 3′ boundary of IVS-C (129 base pairs) shows some differences. A sequence of 185 amino acids deduced from the cloned DNAs represents about 40% of a phaseolin polypeptide.

PHASEOLIN, the first globulin (Gl) fraction to precipitate on dilution of an acidic saline extract[1], is the major storage protein in the seeds of French bean (*Phaseolus vulgaris* L.). One-dimensional SDS-polyacrylamide gel electrophoresis of phaseolin from cotyledons of the cultivar 'Tendergreen' resolves three polypeptides, α, β, and γ, with apparent molecular weights of 53,000, 47,000 and 43,000, respectively[2]. Analytical ultracentrifugation reveals the free polypeptide (3S) conformation in highly alkaline (>pH 11) solutions (or in solution containing SDS), but as the pH values are lowered (in the

---

[†] Present addresses: Atlantic Richfield Co., 6905 Sierra Court, PO Box 2600, Dublin, California 94566 (S.M.S.); Departments of Chemistry and Biochemistry, Southern Illinois University, Carbondale, Illinois 62901 (J.L.S.).

[‖] To whom correspondence should be addressed.

0028-0836/81/020037—05$01.00

absence of dissociating agents), reversible association to 7.1S protomeric (at neutral *p*H values) and 18.2S tetrameric (at *p*H 3.5) forms occurs[3-5]. Mapping according to the method of Cleveland *et al.*[6] established that the α, β and γ polypeptides share a general sequence homology for some 60% of their length[7]. However, two-dimensional electrophoresis[8] shows the presence of charge microheterogeneity[9]. It is likely, therefore, that phaseolin is encoded by a multigene family, of an as yet undetermined size. Similar, or even more complex, situations are indicated for other seed crop proteins such as zein[10-12] from corn and the soybean storage proteins[13,14].

Phaseolin accumulates rapidly in the developing seed cotyledon over a 2- to 3-week period following anthesis[15,16]. Indeed, we estimated that, at certain stages of development, phaseolin polypeptides represent some 90% of the total protein synthetic activity of this tissue. At maturity it comprises about 50% of the total protein stored in the seed[17]. Phaseolin does not occur anywhere in the bean other than in the cotyledon[15]; hence, its synthesis is under strict tissue and temporal regulation.

We describe here the isolation and characterization of a phaseolin gene. This is a major step towards investigating the expression of this economically important protein at the molecular level, and subsequently, in the use of molecular biological techniques to improve the bean as a crop plant[18].

Preparation of bean DNA and construction of a genomic library

The preparation of a genomic library[19-21] of bean DNA was particularly attractive because it makes possible the isolation of phaseolin and other genes of specific interest so that their structure, organization and regulation can be studied. It also facilitates determination of the number of genes coding for phaseolin polypeptides, their variability and cellular location. Only limited information concerning these questions can be obtained by classical genetic techniques. An initial step in preparing such a library was the isolation of DNA of adequate length, free from contaminating polysacharides or RNA, yet without nicks. We could prepare DNA of over 60 kilobase pairs in length from embryonic axes dissected from dry bean seeds (see Fig. 1 legend and ref. 22). After partial digestion of the DNA with *Eco*RI endonuclease, DNA fragments of 8–20 kilobase pairs were selected[23] and inserted into the phage λ cloning vector, Charon 24A. Approximately 2×10^6 recombinant phages were screened for phaseolin sequences using a ^{32}P-cDNA probe made from highly purified phaseolin mRNA[23,24]. From 12 candidate clones, two, Ch24A-176.2 and Ch24A-177.4, were selected for further characterization.

Characterization of cloned DNA

Restriction analysis of 'mini-lysates'[19] of clones Ch24A-176.2 and Ch24A-177.4 showed that they both contained a 7.2.-kilobase *Eco*RI fragment (Fig. 1, lanes *b*, *c*) that hybridized strongly the cDNA probe (Fig. 1, lanes *d*, *e*). DNA from clone Ch24A-177.4 was purified[25] for further analysis. This revealed (Fig. 1, lanes *f*, *g*) that, in addition to a strongly hybridizing 7.2-kilobase fragment, weak hybridization to a 1.7-kilobase fragment could also be discerned. An additional 2.2-kilobase fragment was detected that showed no hybrization to the probe. To facilitate the preparation of large quantities of cloned DNA, the 7.2- and 1.7-kilobase DNA fragments were subcloned into pBR322 (refs 20, 26).

Although the material used as a template for the synthesis of the ^{32}P-cDNA probe was known to be highly enriched in phaseolin mRNA[23,27], it was necessary to confirm that the cloned DNA did contain a phaseolin gene sequence. The first confirmation was obtained by electrophoretic separation of bean 18S and 28S rRNAs, poly(A)$^-$RNA, total poly(A)$^+$RNA, and the 16S fraction of poly(A)$^+$RNA (phaseolin mRNA) in an agarose gel (Fig. 2*A*). After electrophoresis, the RNAs were transferred to diazobenyloxymethyl (DBM)-paper[28,29] and then hybridized to the 7.2-kilobase cloned DNA fragment which was ^{32}P-labelled by nick-translation[30]. No hybridization to rRNA

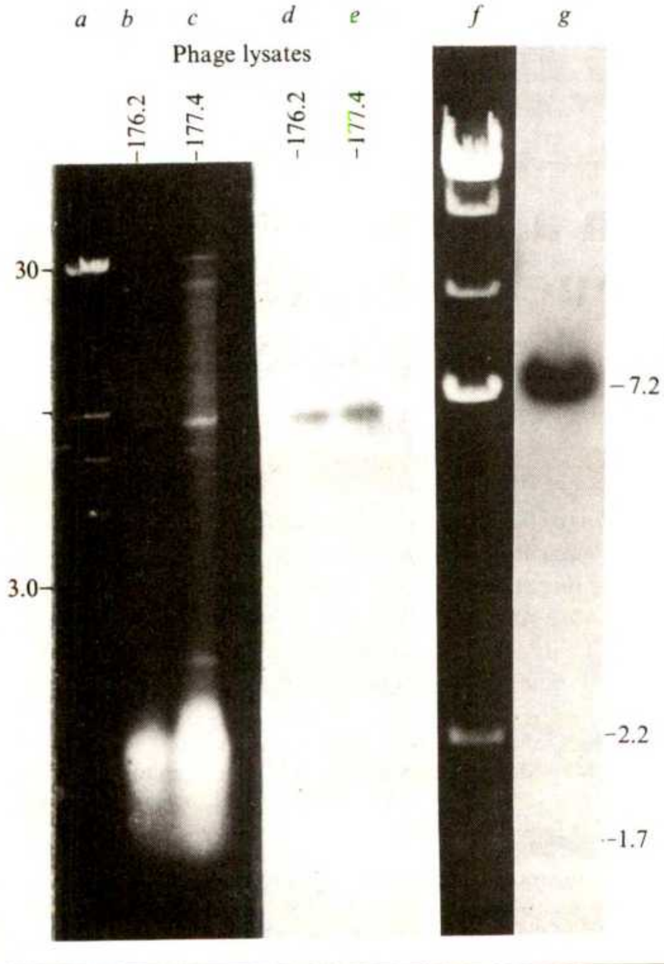


Fig. 1 Identification and selection of clones containing phaseolin sequences from a bean DNA library. French bean genomic DNA was isolated from embryo axes of dry seeds by the method of Blin and Stafford[22] with the following modifications. After extraction with phenol and dialysis, the DNA solution was adjusted to 0.95 g ml^{-1} CsCl and 0.5 mg ml^{-1} ethidium bromide, then centrifuged twice for 60 h (45,000 r.p.m., Beckman Ti-60) at 20 °C to band the DNA, which was recovered from the gradient after each cycle. Ethidium bromide was removed by several extractions with CsCl-saturated isopropyl alcohol and the DNA dialysed against 10 mM Tris-Cl, *p*H 8.0, 5 mM NaCl and 1 mM EDTA before use. The DNA obtained had a length of about 60 kilobase pairs. An additional extraction with phenol was necessary after the dialysis step to obtain normal *Eco*RI digestion kinetics and higher cloning efficiencies. The construction of a 'shotgun' collection of 1.5×10^6 Charon 24A bacteriophage recombinants containing partial *Eco*RI fragments from 10 µg of genomic DNA was as described by Blattner *et al.*[19] and Slightom *et al.*[20]. This original library was amplified to 2×10^9 recombinant phages and approximately 2×10^6 of these were screened[19] using a ^{32}P-labelled phaseolin cDNA probe prepared according to Kacian and Myers[24]. DNA was extracted from mini-lysates[19] of two clones that showed especially strong plaque hybridization to the probe. *Eco*RI digests of these clones were electrophoresed in a 0.7% agarose gel (lane *b*, clone 176.2; lane *c*, clone 177.4; lane *a* contained marker DNAs[19] whose sizes are shown at the left in kilobases). The DNA was then transferred to nitrocellulose paper by the method of Southern[40] and hybridized to the phaseolin ^{32}P-cDNA probe. The 7.2-kilobase *Eco*RI fragment from each clone hybridized strongly to the probe (lanes *d*, *e*). After purification of DNA from clone Ch24A-177.4 and digestion with *Eco*RI, two additional bands (2.2 and 1.7 kilobases, see lane *f*) were revealed. The ^{32}P-cDNA probe hybridized weakly to the 1.7-kilobase fragment but showed no recognition of the 2.2-kilobase fragment (lane *g*).

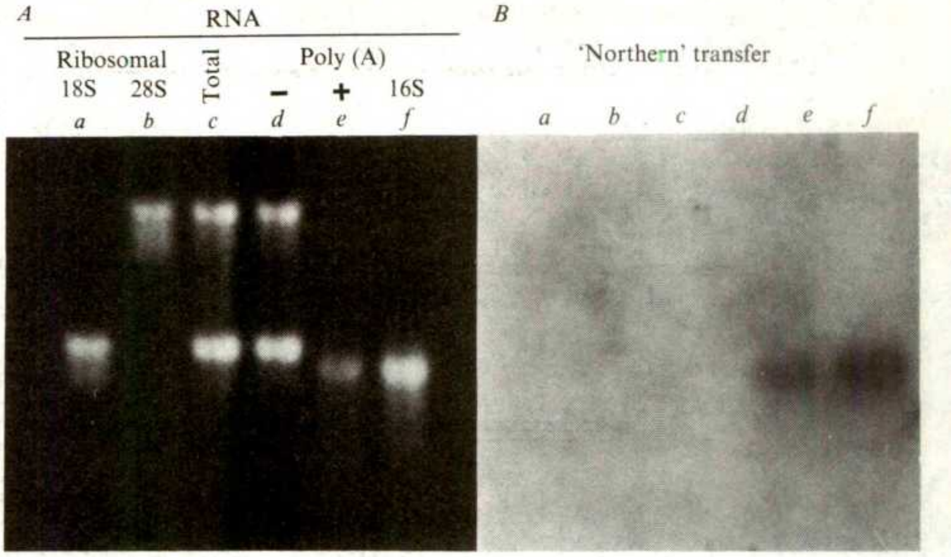


Fig. 2 Hybridization of cloned DNA to bean RNAs. Total RNA was isolated from maturing French bean cotyledons (grown in the Madison campus Biotron) using a hot (~70 °C) Na borate-NaOH (0.2 M, *p*H 9.0) extraction buffer containing 1% SDS, 30 mM EGTA and 5 mM dithiothreitol, followed by deproteinization at 37 °C with proteinase K (Beckman) as described by Buchbinder[41] and Hall *et al.*[23,27]. The total RNA was separated into poly(A)$^+$ and poly(A)$^-$ fractions by oligo(dT)-cellulose chromatography. The poly(A)$^+$RNA was fractionated by centrifugation in a linear–log[42] sucrose gradient (0–32.5% sucrose in 0.1 M NaOAc, 0.1 M NaCl and 1 mM EDTA, *p*H 5.0), and the 16S fraction collected. These RNA fractions, together with 18S and 28S bean rRNAs, were electrophoretically separated on a 1.5% agarose gel containing 6% formaldehyde and stained with ethidium bromide (*A*: lane *a*, 18S rRNA; lane *b*, 28S rRNA; lane *c*, total RNA; lane *d*, poly (A)$^-$ RNA; lane *e*, poly(A)$^+$RNA; lane *f*, 16S RNA), and subsequently transferred to DBM-paper (Schleicher and Schuell) according to the procedures of Alwine *et al.*[28] and Rave *et al.*[29]. The RNAs bound to the paper were then hybridized to the cloned 7.2.-kilobase bean DNA fragment labelled with ^{32}P by nick translation[30]. As seen in *B*, the radioactive cloned DNA hybridized to only the poly(A)$^+$ (lane *e*) and 16S (lane *f*) RNAs.

was seen, only the total poly(A)$^+$ and the phaseolin mRNAs showed positive reactions (Fig. 2*B*). This established that the clone does not contain any rRNA sequences and shows that it reacts with the expected poly(A)$^+$RNA fractions.

A second confirmation came from hybrid-arrested translation[31], in which only those mRNAs complementary to the supplied DNA sequence are prevented from functioning as messengers during cell-free protein synthesis. Figure 3*A* shows that increasing amounts of cloned DNA competed out products identified as phaseolin polypeptides, although synthesis of the other products coded by the bean cotyledon poly(A)$^+$RNA was unaffected. Further confirmation that a phaseolin gene had been isolated was obtained by binding the recombinant DNA to a nitrocellulose filter and hybrid-selecting[32] phaseolin mRNA

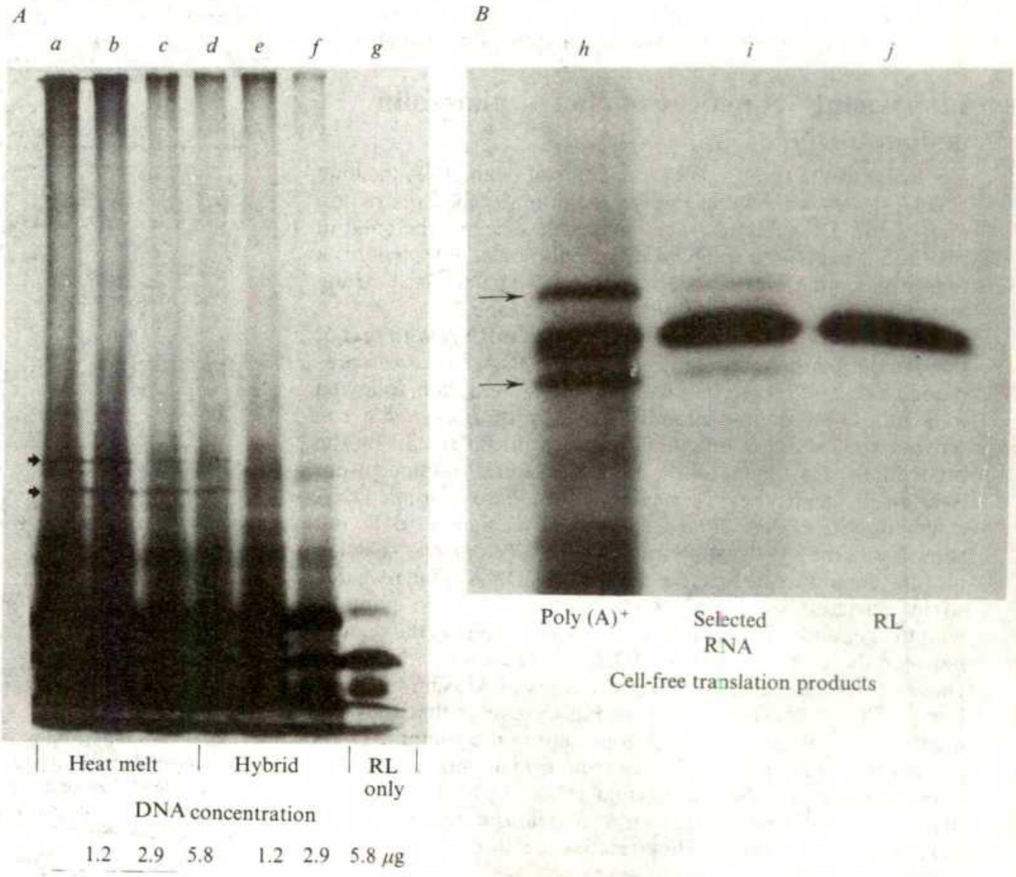


Fig. 3 Hybrid-arrested and hybrid-selected translation of phaseolin mRNA. *A*, Hybrid-arrested translation[31] in the rabbit reticulocyte system was performed using various amounts (2.4 μg, 5.8 μg or 11.6 μg) of *Eco*RI-digested Ch24A-177.4 DNA hybridized to 0.4 μg bean poly(A)$^+$RNA for 2 h at 48 °C in the presence of 80% deionized formamide. The samples (250 μl) were each divided into two equal portions, one of which was heat melted at 100 °C for 60 s before translation (lanes *a*, *b*, *c*; heat melt) and the other allowed to remain in hybrid form during translation (lanes *d*, *e*, *f*; hybrid). A reaction containing no added mRNA is shown for comparison in lane *g* (RL only). The ^{35}S-Met-labelled products were separated on a 13% SDS-polyacrylamide slab gel[43] and subjected to fluorography[44]. The arrows indicate the position of phaseolin polypeptides. *B*, Phaseolin mRNA was selected from total poly(A)$^+$mRNA by hybridization with pBR322-7.2 kilobase pair DNA (plasmid containing the 7.2-kilo base pair genomic fragment), immobilized on a nitrocellulose filter according to the procedures of McKnight *et al.*[32]. DNA (32.5 μg) was digested with *Eco*RI, the protein extracted with phenol and then precipitated with ethanol before immobilization. Hybridization was carried out in PIPES buffer (950 mM, *p*H 7.5) containing 50 μg ml^{-1} bean RNA, 2 mM EDTA, 0.4% SDS, 0.5 M NaCl and 30% formamide at 47 °C for 2 h with gentle shaking. After being thoroughly washed, the hybridized RNA was eluted from the filter and translated. The products of cell-free translation were analysed as described for *A*. Lane *h*, products coded by bean poly(A)$^+$ RNA; lane *i*, products coded by RNA selected by cloned phaseolin DNA; lane *j*, products made in the absence of added RNA, showing the prominent band (apparent 44,000) observed when ^{35}S-Met is used as substrate in the reticulocyte system.

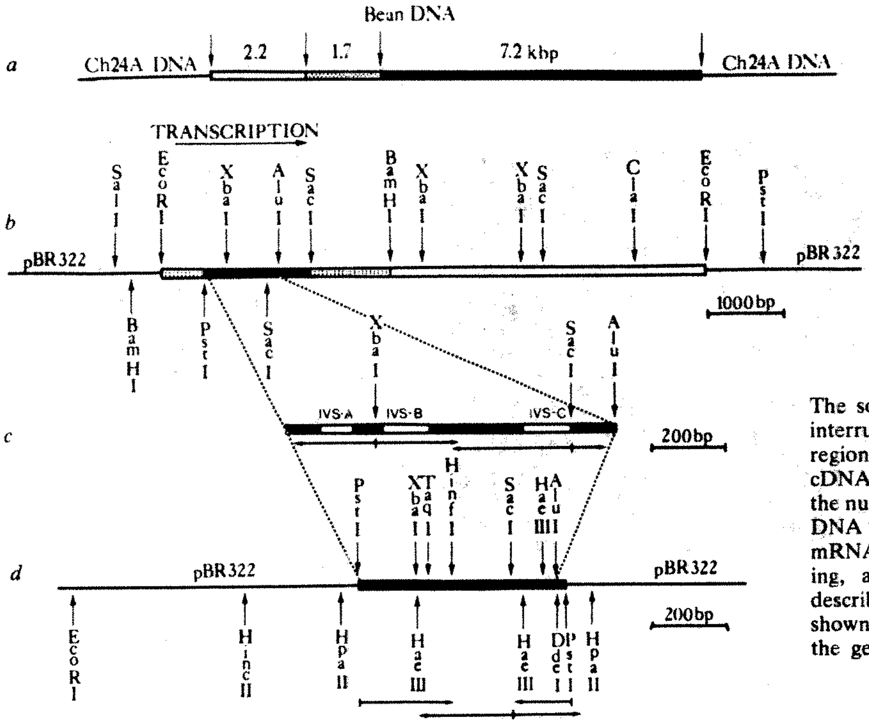


Fig. 4 Restriction maps of bean genomic DNA and phaseolin cDNA clones. *a*, Arrangement of bean DNA in clone Ch24A-177.4. Vertical arrows indicate *Eco*RI sites. The solid black line shows the region of the bean DNA that hybridizes strongly to phaseolin cDNA, the stippled area denotes weak hybridization and the white area indicates no hybridization. *b*, Detailed map of the 7.2-kilobase *Eco*RI fragment of bean DNA after subcloning in pBR322. The direction of transcription shown above frame *b* was oriented by the GT-AG rule of intervening sequences[36] and by the method described by Fryberg *et al.*[45]. *c*, Sequenced portion of the 7.2-kilobase DNA (see Fig. 5). The solid black areas indicate coding regions that are interrupted by three intervening sequences (white regions), IVS-A, IVS-B and IVS-C. *d*, Map of phaseolin cDNA in clone pBR322.74. The solid black area contains the nucleotide sequence shown in Fig. 5. Complementary DNA was prepared by reverse transcription of 16S bean mRNA[24]; synthesis of double-stranded cDNA, G-C tailing, and cloning into pBR322 were performed as described by Weinand *et al.*[33]. The horizontal arrows shown in *c* and *d* indicate the direction of sequencing for the genomic and cDNA clones, respectively. bp, Base pairs.

from a total poly(A)$^+$RNA preparation (Fig. 3*B*). Cell-free translation of the hybridized mRNA showed that only mRNA coding for phaseolin proteins had been removed by the genomic clone (Fig. 3*b*, lane *i*); no additional protein bands were observed. In agreement with the extensive sequence homology revealed by peptide mapping[7], the messenger-arrest and messenger-selection procedures affected translation of both phaseolin polypeptides *in vitro* (processing events, including glycosylation, give rise to the three polypeptides resolved for the native protein[23]).

Intervening sequences in cloned phaseolin genomic DNA

The order of the three *Eco*RI fragments of bean DNA in clone Ch24A-177.4 is shown in Fig. 4*a*. A more detailed restriction map of the 7.2-kilobase fragment that has been subcloned in pBR322 is presented in Fig. 4*b*. The solid black line denotes a region of about 1,300 base pairs that showed particularly strong hybridization to the phaseolin ^{32}P-cDNA probe.

To compare the squence of the phaseolin mRNA with that of the cloned phaseolin genomic DNA, a cDNA clone was prepared by reverse transcription of the mRNA template, followed by synthesis of double-stranded DNA with *Escherichia coli* polymerase I, G-C tailing, and insertion into pBR322 by the procedures of Weinand *et al.*[33]. cDNA clones obtained were hybridized against the ^{32}P-labelled 7.2-kilobase genomic DNA fragment; 30 cDNA clones gave positive signals (data not shown) and the strongest hybridizing cDNA clone was selected for further study. A restriction map of this cDNA clone revealed some restriction sites (*Pst*I, *Xba*I, *Sac*I and *Alu*I) in common with the genomic clone (compare Fig. 4*b*, *d*). Hence, the regions between these sites in both the cDNA and genomic clones were chosen for sequencing by the procedures of Maxam and Gilbert[34]. The comparison shown in Fig. 5 reveals that they are identical except for three regions present in the genomic DNA but absent from the cDNA. These sequences are presumed to be intervening sequences (introns) that are deleted during processing of the initial nuclear transcript to yield the mature mRNA found in the cytoplasm[35]. Their relative locations are mapped in Fig. 4*c*.

All three intervening sequences (introns) of this plant gene begin with the nucleotides GT and end with AG, as do those of animal and virus genes[36]. Recently, Lerner *et al.*[37] have compared the regions bounding introns of some 43 eukaryotic DNAs. Their study revealed a general pattern for these regions,

```
   ↓5' END OF cDNA CLONE 25                       50                       75
g ...........................................................................
c GAACCATTTTTCTATTTGGTTAACCCTGATCCCAAAGAGGATCTCAGAATAATCCAACTCGCCATGCCCGTTAAC
a  E  P  F  F  Y  L  V  N  P  D  P  K  E  D  L  R  I  I  Q  L  A  M  P  V  N
                        100                      125                      150
g ...............GTACTGCCTTTTGTAATACCGAACTAATTTTTTGTTATTTTAACTTGCAATTTCTCTCCA
c AACCCTCAGATTCAT************************************************************
a  N  P  Q  I  H|←----------------------- IVS-A ----------------------------
                        175                      200                      225
g AATGTGATGATAAATGTTTGTCCTGTAG...............................................
c ****************************GAATTTTTCCTATCTAGCACAGAAGCCCAACAATCCTACTTGCAAGA
a -------- IVS-A ----------->|E  F  F  L  S  S  T  E  A  Q  Q  S  Y  L  Q  E
                XbaI    250                      275                      300
g ..................................GTAAGAAAGAAAATAGTATCTAACTATATATTTGCGTTGCT
c GTTCAGCAAGCATATTCTAGAGGCCTCCTTCAAT*****************************************
a   F  S  K  H  I  L  E  A  S  F  N|←-------------- IVS-B -----------------
                        325                      350                      375
g ATTTAGCTAGTACCTTGTCCAAATGTCACACCTGTTGAATTTGTTGAATGATATCATTATATATGTTTGCATGAT
c ***************************************************************************
a -------------------------------- IVS-B -----------------------------------
                TaqI    400                      425                      450
g TTTTATAG...................................................................
c ********AGCAAATTCGAGGAGATCAACAGGGTTCCGTTTGAAGAGGAGGGACAGCAAGAGGGAGTGATTGTGA
a IVS-B →|S  K  F  E  E  I  N  R  V  P  F  E  E  E  G  Q  Q  E  G  V  I  V
       HinfI            475                      500                      525
g ...........................................................................
c ACATTGATTCCGAACAGATTAAGGAACAGAGCAAACATGCAAAATCTAGTTCAAGGAAATCCCCTTCCAAACAAG
a N  I  D  S  E  Q  I  K  E  Q  S  K  H  A  K  S  S  S  R  K  S  P  S  K  Q
                        550                      575                      600
g ...........................................................................
c ATAACACAATTGGAAACGAATTTGGAAACCTGACTGAGAGGACCGGTAACTCCCTGAATGTGTTAATCAGTTCCA
a D  N  T  I  G  N  E  F  G  N  L  T  E  R  T  G  N  S  L  N  V  L  I  S  S
                        625                      650                      675
g ..............GTAAATACAAGGAAAAACCATATAGACAAATCTAGCAATTAGATTCTATTATTCGACTGTC
c TAGAGATGGAAGAG*************************************************************
a I  E  M  E  E|←--------------------------- IVS-C -------------------------
                        700                      725                  SacI750
g GTCTTGGTTAGAAAATCTTAGTATTGAGACTGTGGTTGGGTGGTGGTTTTTTTTGTTGGCGGATTTAG.......
c ********************************************************************GGAGCTC
a ------------------------ IVS-C ----------------------------------->|G  A
                        775HaeIII                800                      825
g ...........................................................................
c TTTTTGTGCCACACTACTATTCTAAGGCCATTGTTATACTAGTGGTTAATGAAGGAGAAGCACATGTTGAACTTG
a L  F  V  P  H  Y  Y  S  K  A  I  V  I  L  V  V  N  E  G  E  A  H  V  E  L
                        850           AluI       DdeI                     900
g .......................................↓3' END OF GENOMIC CLONE
c TTGGCCCAAAAGGAAATAAGGAAACCTTGGAATATGAGAGCTACAGAGCTGAGCTTTCTAAAGACGATGTCCCCC→
a V  G  P  K  G  N  K  E  T  L  E  Y  E  S  Y  R  A  E  L  S  K  D  D  V|pBR322
                                                       3' END OF cDNA CLONE↑
```

Fig. 5 Comparison of phaseolin genomic and cDNA sequences and the deduced polypeptide sequence. The upper line (*g*) gives the sequence of the portion of genomic DNA denoted in Fig. 4*c*. Dots indicate that the genomic sequence is identical to that of the phaseolin cDNA clone shown in the middle line (*c*). Asterisks denote intervening sequences (*IVS-A*, 88 base pairs; *IVS-B*, 124 base pairs; *IVS-C*, 129 base pairs) that are present in the genomic clone but absent from the cDNA clone. The sequence of 185 amino acids corresponding to the cDNA sequence is shown in the bottom line (*a*) by the single-letter code[46]. ^{32}P-end labelling and DNA sequencing were according to the procedure of Maxam and Gilbert[34], as modified by Slightom *et al.*[20].

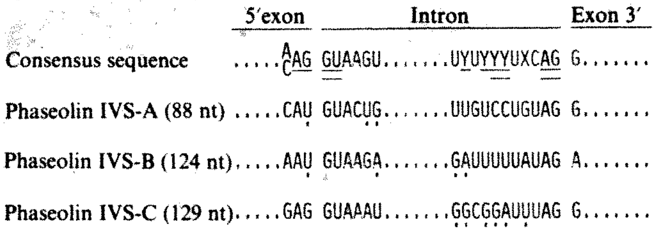

5′exon
Intron
Exon 3′
Consensus sequence .....A/C AG GUAAGU.......UYUYYYUXCAG G.......
Phaseolin IVS-A (88 nt) .....CAU GUACUG.......UUGUCCUGUAG G.......
Phaseolin IVS-B (124 nt).....AAU GUAAGA.......GAUUUUUAUAG A.......
Phaseolin IVS-C (129 nt).....GAG GUAAAU.......GGCGGAUUUAG G.......

Fig. 6 Comparison of phaseolin intervening sequences with that of a consensus sequence. The consensus sequence shown is from Lerner *et al.*[37]; single underlines indicate a frequency of occurrence of 45% or greater (in 43 sequences) and double underlines a frequency of 95% or more. Dots under nucleotides (nt) denote differences from the consensus sequence where a purine, X (A or G) replaces a pyrimidine, Y (U or C).

suggesting an involvement in recognition by enzymes responsible for intron excision and exon splicing[38]. A 'consensus' sequence was derived, and this is compared with the intron–exon boundaries of the phaseolin gene in Fig. 6. For phaseolin introns IVS-A and IVS-B, the boundaries agree closely with the consensus sequence, but 6 of the 11 nucleotides at the 3′-end of IVS-C differ from the generalized sequence. Of these six positions, two purine residues replace pyrimidine locations found in 95% of the sequences analysed by Lerner *et al.*[37].

Amino acid sequence from the cloned nucleotide sequences

The DNA sequences were subjected to computer analysis to determine the derived amino acid sequence. Only one open reading frame was found in the DNA sequence of the cDNA clone, and the corresponding amino acid sequence (representing a fragment of about 40% of the length of a phaseolin polypeptide) is shown in Fig. 5. The amino acid composition of this fragment seems to be representative of the overall protein as a good agreement (coefficient of correlation 0.95) was obtained between the derived sequence and that found by amino acid analysis of the native protein[39].

Peptide mapping of phaseolin positioned two methionine residues some 12 kilodaltons apart, with one residue located about 6 kilodaltons from one end of the polypeptide[7]. Only two methionine residues were present in the derived sequence shown in Fig. 5 and these are 110 amino acid residues (about 12 kilodaltons) apart. It seems likely, therefore, that the derived sequence starts some 6 kilodaltons from the N-terminus of the phaseolin polypeptide. The peptide mapping data[7] also showed that the phaseolin polypeptides have an essentially identical amino acid sequence for some 60% of their length; the heterologous region lying to the C-terminal (direction deduced from the present data) end of the molecules is presumed to be primarily responsible for the multiple phaseolin polypeptides resolved in high-resolution two-dimensional electrophoresis[9]. The cloned DNA fragments are located within the homologous region of phaseolin. Therefore, the correspondence of the genomic and cDNA sequences, and also the cross-hybridization to multiple phaseolin mRNAs (as evidenced by the hybrid-selection and hybrid-arrest data), are not surprising.

Interestingly the derived sequence reveals that 13 of the 19 aspartate residues are asparagine and of the 36 glutamate residues, 10 are glutamine; that is, there are 23 amide residues in the defined sequence of 185 amino acids. The high amide content of phaseolin reflects its role as a N-storage protein for the seed, but a high proportion of acidic residues is also expected from the general properties of this protein.

The sequence CG is rarely found in either the first or second (absent, but expected four times by random chance) or second and third (occurs once; random expectation is seven times) codon positions. Thus, of the six arginine codons available, only two are used to an appreciable extent in the sequence analysed thus far.

Conclusions

From the data available, most of the region coding for phaseolin probably lies within the area denoted by the heavy black line in Fig. 4*a*, *b*. The remainder is presumed to be within the stippled regions at each side of this area because they show weaker hybridization to the phaseolin probe and because no hybridization occurs outside these regions. Thus, it seems that clone Ch24A-177.4 contains an entire phaseolin gene, together with extensive sequences flanking its 3′ and 5′ ends.

Our results show that the structure of this plant gene is similar to that of animal DNAs in containing intervening sequences. This isolation of one of the genes encoding the major storage globulin in bean seeds should provide a basis for exploring regulatory constraints for the expression of this nutritionally significant protein.

We thank Dr O. Smithies (NIH grant GM-20069) and his associates, especially Tom Szeto, Ann Blechl and Natalie Borenstein, for guidance in cloning operations and for the use of his P2 facility in which all recombinant operations were carried out according to current NIH guidelines. We also thank Dr W. M. Fitch for the computer derivation of the amino acid code from the nucleotide sequence and Dr J. Ross for advice and a gift of S_1 nuclease. Reverse transcriptase was obtained from Dr J. W. Beard of Life Sciences through the Viral Cancer Program of the NCI and the Charon 24A was a gift from F. R. Blattner. This study was supported by grants from USDA/SEA (5901-0410-9-0357-0), NSF (PCM 78-11804), the Herman Frash Foundation and by the Research Division of the College of Agricultural and Life Sciences (Hatch 1482) of the UW-Madison.

Received 25 July; accepted 24 October 1980.

1. Osborne, T. B. *J. Am. chem. Soc.* **16,** 633–764 (1894).
2. McLeester, R. C., Hall, T. C., Sun, S. M. & Bliss, F. A. *Phytochemistry* **12,** 85–93 (1973).
3. Sun, S. M., McLeester, R. C., Bliss, F. A. & Hall, T. C. *J. biol. Chem.* **249,** 2118–2121 (1974).
4. Pusztai, A. & Watt, W. B. *Biochim. biophys. Acta* **207,** 413–431 (1970).
5. Joubert, F. J. *J. S. Afr. Chem. Inst.* **10,** 16–20 (1957).
6. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. L. *J. biol. Chem.* **252,** 1102–1106 (1977).
7. Ma, Y., Bliss, F. A. & Hall, T. C. *Pl. Physiol.* **66,** 897–902 (1980).
8. O'Farrell, P. H. *J. biol. Chem.* **250,** 4007–4021 (1975).
9. Brown, J. W. S., Bliss, F. A. & Hall, T. C. *Pl. Physiol.* **6,** 838–840 (1980).
10. Viotti, A. *et al. Eur. J. Biochem.* **102,** 211–222 (1979).
11. Park, W. D., Lewis, E. D. & Rubenstein, I. *Pl. Physiol.* **65,** 98–106 (1980).
12. Larkins, B. A. *et al.* in *Genome Organization and Expression in Plants* (ed. Leaver, C. J.) 203–217 (Plenum, New York, 1980).
13. Beachy, R. N., Barton, K. A., Madison, J. T., Thompson, J. F. & Jarvis, N. in *Genome Organization and Expression in Plants* (ed. Leaver, C. J.) 273–281 (Plenum, New York, 1980).
14. Moreira, M. A., Hermodson, M. A., Larkins, B. A. & Nielsen, N. C. *J. biol. Chem.* **254,** 9921–9926 (1979).
15. Sun, S. M., Mutschler, M. A., Bliss, F. A. & Hall, T. C. *Pl. Physiol.* **61,** 918–923 (1978).
16. Mutschler, M. A., Bliss, F. A. & Hall, T. C. *Pl. Physiol.* **65,** 627–630 (1980).
17. Hall, T. C. *et al.* in *The Plant Seed* (eds Rubenstein, I., Phillips, R. L., Green, C. E. & Gengenbach, B. G.) 3–25 (Academic, New York, 1979).
18. Bliss, F. A. & Hall, T. C. *Cereal Foods World* **22,** 106–113 (1977).
19. Blattner, F. R. *et al. Science* **202,** 1279–1284 (1978).
20. Slightom, J. L., Blechl, A. E. & Smithies, O. *Cell* **21,** 627–638 (1980).
21. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
22. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3,** 2303–2308 (1976).
23. Hall, T. C. *et al.* in *Genome Organization and Expression in Plants* (ed. Leaver, C. J.) 259-272 (Plenum, New York, 1980).
24. Kacian, D. L. & Myers, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2191–2195 (1976).
25. Moore, D. D., Denniston-Thompson, K., Furth, M. E., Williams, B. G. & Blattner, F. R. *Science* **198,** 1041–1046 (1977).
26. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
27. Hall, T. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3196–3200 (1978).
28. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
29. Rave, N., Crkvenjakov, R. & Boedtker, H. *Nucleic Acids Res.* **6,** 3559–3567 (1979).
30. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
31. Paterson, B. M., Roberts, B. Z. & Kuff, E. L. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4370–4374 (1977).
32. McKnight, G. S., Lee, D. C., Mayo, K. E. & Palmiter, R. *Nucleic Acids. Res.* **6,** 2435–2452 (1979).
33. Wienand, U., Brüschke, C. & Feix, G. *Nucleic Acids Res.* **6,** 2707–2715 (1979).
34. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–565 (1977).
35. Tilghman, S. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 725–729 (1978).
36. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4853–4857 (1978).
37. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283,** 220–224 (1980).
38. Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1877–1879 (1980).
39. Sun, S. M. thesis, Univ. Wisconsin, Madison (1974).
40. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
41. Buchbinder, B. U. thesis, Univ. Wisconsin, Madison (1980).
42. Brakke, M. K. & Van Pelt, N. *Analyt. Biochem.* **38,** 56–64 (1970).
43. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
44. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
45. Fyrberg, E. A., Kindle, K. L., Davidson, N. & Sodja, A. *Cell* **19,** 365–378 (1980).
46. IUPAC-IUB Commission *Biochem. J.* **113,** 1–4 (1969).

LETTERS

Spectral evolution of the 5 March 1979 γ burst

E. E. Fenimore, W. D. Evans, R. W. Klebesadel, J. G. Laros & J. Terrell

University of California, Los Alamos Scientific Laboratory,
Los Alamos, New Mexico 87545

The γ burst[1] of 5 March 1979 was observed by nine experiments[2-6] widely spaced on an interplanetary scale allowing an accurate position for a γ-burst source to be determined[2,6] for the first time. Spectral observations in the range 50–1,000 keV had sufficient time resolution to resolve spectral changes within the main burst. Early in the main burst, the typical photon energies declined from a spectrum similar to a 30 keV black body to that of a ~26 keV black body. Thereafter, at the shoulder on the main burst the spectrum became harder ($KT \sim 31$ keV). The γ burst apparently triggered an X-ray pulsar with a period of ~8.0 s (refs 3–5, 7) and with a softer spectrum ($KT \simeq 12$ keV). We show here that, although the source is probably a neutron star with an obliquely rotating magnetic field, simple cooling of residual hot spots does not completely explain the pulsations.

Figure 1 presents the time history and spectral development of the 5 March 1979 γ-burst as observed by the Pioneer Venus Orbitor (PVO) detector (see ref. 8 for instrumental details). The PVO instrument varies its sampling rate depending on the count rate, which was near saturation during the burst. As a result, the intensity and spectra were recorded at the highest sampling rate of any of the satellites observing the event (either every 0.25 ms or 0.5 ms for the intensity and every ~6 ms for the spectra). Figure 1*a*, *b* presents the intensity and a hardness ratio for the burst phase of the event, Fig. 1*c*, *d* shows the intensity and hardness ratio plotted on a scale intended to display the 8-s periodicity[3-5,7]. The intensities and the hardness ratios in Fig. 1 have been binned to improve the statistics.

The hardness ratio is defined to be the ratio of the sum of (background-corrected) counts in three channels spanning 100–1,000 keV to the counts in the 50–100 keV channel. Each ratio is represented by a box whose width is the range of time of the sample and whose height is determined by $\pm 1.0\sigma$ error limits. By folding a range of black-body spectra with the instrumental response, this ratio has been mapped into an effective black-body temperature at the right-hand side of Fig. 1*d*. A black-body form was used to parameterize the data in a manner which allows an estimate of the size of the emitting region by Stefan's law.

A key feature of this burst is its very fast rise time. A reanalysis[7] of the PVO data reveals that rise times of ≤ 1 ms are consistent with the data. (This value is larger than our original[2] value of 0.25 ms.) After the sharp rise, the intensity continues to rise for ~25 ms. During this rise and until $t = 75$ ms the hardness ratio has a downward trend. Black bodies with KT from 30 to 26 keV can reproduce the observed hardness ratio during this phase. However, the observed distribution of counts has too many counts in the two higher energy bins (250–500 and 500–1,000 keV) to be entirely consistent with black-body radiation. Thus, during this phase, the hardness ratio and the effective kT should be used only as an indicator of typical energies involved. Starting at ~80 ms after the initial rise, there is a shoulder on the light curve of the main burst which lasted for ~0.1 s. This shoulder has a tendency to have a somewhat harder spectrum (~31 keV).

The significance of these variations was checked by χ^2 tests. Two ranges were tested, from $t = 0.0$ to 75 ms and from $t = 0.0$ to 120 ms. Random fluctuations could account for the observed distribution of hardness ratios only 3% and 2% of the time, respectively. We conclude that there is a statistically significant trend for the spectrum to soften initially (until t ~75 ms) and then harden during the shoulder on the γ-burst light curve. In contrast to the first 80 ms, the spectrum during the shoulder can be fit by a black-body spectrum (among other spectral forms) to within the accuracy of the statistics. As the intensity within the shoulder decays to near background (~150 c.p.s. (counts per s) at t~2.0 s), the spectrum softens dramatically.

Following the γ-burst, the source exhibited many of the characteristics of a hard X-ray pulsar[3]. The period was observed to be 8.0 ± 0.1 (refs 3–5, 7). Figure 1*c* shows the time history for the first few cycles of the pulsations. (For longer time histories where the periodicity is more evident see refs 3–5, 7). Note that, particularly for the first pulse, the hardness ratio varies in phase with the intensity from near zero between the pulses to a KT_{eff} of ~12 keV at the peak of the pulse (see Fig. 1*d*). This statistically significant variation in phase with the intensity is probably due to a varying look angle to the source as the object rotates. Due to deteriorating statistics and a greater contribution of the interpulse, it is not possible to determine whether similar spectral changes occur in the subsequent pulses.

The short rise time, the pulse period, the regularity of the pulses, the existence of an interpulse, and the observation of a line at a location consistent with the 511-keV line[3] redshifted by a 1 solar-mass neutron star, all argue that the pulsations are due to hot spots at the magnetic poles of a rotating neutron star. If so, a possible explanation for the difference between the γ-burst temperature and the X-ray pulsar temperature is that the X-ray pulsar is due to residual hot spots at the poles which are cooling.

If the observed lower temperature ($KT \sim 12$ keV) in the pulses represents a cooling from the initial hotter γ-burst, one would expect the side of the star giving the γ-burst to show a continually decreasing temperature and count rate. Figure 2

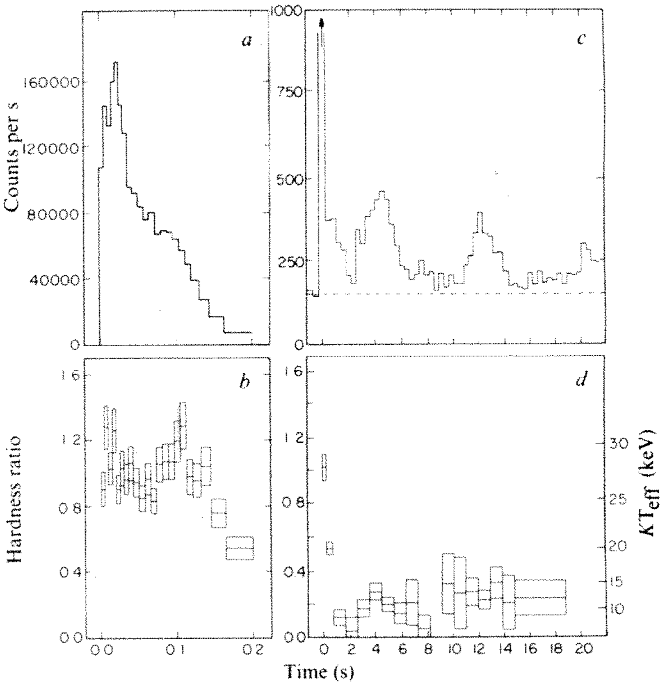


Fig. 1 The Pioneer Venus Orbiter time history and hardness ratio for the 5 March 1979 γ-burst. Initial time t_0 is at 57118.019 s UT. The dotted line in *c* is the background level. *a*, *b*, Burst; *c*, *d*, 8-s pulsation.

0028-0836/81/020042—02$01.00

contrasts the behaviour of the two sides of the star and is based on the Venera 11 and 12 observations[3] because they had the highest sensitivity and were made from three-axis stabilized satellites. Each point on the histogram represents the average intensity (scaled to the PVO sensitivity) near the peak of a pulse, determined by summing the Venera 11 and 12 observations[3] for 2 s. The cross-hatched regions represent the pole on the side of the star from which the γ-burst was observed. The diagonally striped regions are due to the opposite pole.

The flux from the side away from the burst is consistent with cooling (that is, it is decreasing). However, the pole associated with the burst shows an increasing average count rate, being 95 ± 1.57 c.p.s. for the pulse at ~8 s and 113 ± 1.7 for the pulse at $t \sim 16$ s. Applying a standard statistical test, one finds that the difference in count rate represents an $\sim 10\sigma$ deviation. We conclude that, unless the Venera data contains nonstatistical uncertainties, the 2-s averaged count rate from the pole associated with the burst increased between $t \sim 8$ and ~16 s.

Figure 1*d* shows how the spectrum changed between these two intervals. From $t \sim 7.0$ to 8.25 s, the hardness ratio was ~0.05. During the period $t \sim 8.25$–9.25 s, the high energy channels (100–1,000 keV) were below anticipated background, implying a hardness ratio of zero. In contrast to these values, the interval from $t \sim 15.0$ to 18.75 s had a hardness ratio of 0.23. The increase in the observed count rate in the higher energy channels between the two pulses is statistically significant at the 98% confidence level. Also, according to the PVO data the overall observed count rate significantly decreased (5.8σ), implying that the temperature was indeed increasing while either the source size was decreasing or the column density was increasing. However, the decrease in the PVO overall count rate is not in accord with the increase observed from the three-axis stabilized observations of the Veneras. The PVO observations were possibly affected by spin modulation of the satellite. Although the spin axis was pointed to within 3 degrees of the location of the burst[2,6], there could have been sufficient material moving in and out of the field of view to cause the small discrepancy between PVO and the Veneras. In either case (if the intensity increased as observed by the Veneras or if the hardness ratio increased as observed by PVO), the cooling explanation is unlikely. Because the PVO had a possible (although unlikely) source of systematic error, we are inclined to accept the Venera result that the intensity increased. If so, the increase in hardness ratio observed by PVO is less certain.

One possibility consistent with the observations is that the γ-burst altered the magnetosphere, triggering a short-lived (few minutes), accretion-driven, hard X-ray pulsar. The difference in hardness ratio between the γ-burst and the pulsation then is a result of their being produced by separate but related energy sources. One constraint on such a model would be whether the magnetosphere can hold sufficient material without violating the observed upper limit on the steady state X-ray flux[9]. In addition the source would have to be close enough to avoid violating the Eddington limit[10], ruling out N49 as the counterpart. A similar explanation is that the neutron star suffered a collision with a nearly intact comet (as recently proposed in ref. 11 for the 5 March γ-burst) with the pulsations then being due to left over debris accreting down the field lines.

An accreting model for the pulses is not the only possible explanation. For example, if the initial γ rays were collimated into a beam that was not colinear with our line of sight to the pulsar, and if the beam subsequently broadened between $t \sim 8$ s and ~16 s, then the average count rate could increase even if the source were cooling. Indeed, ref. 3 shows that the pulse at $t \sim 8$ s was probably narrower than subsequent pulses, contributing to its lower average count rate. There must be other possible beam geometries that could have a similar effect.

The nature of the γ-burst remains uncertain. We will not consider magnetospheric accretion because current models[12,13] predict rise times that are 100–1,000 times larger than observed for the γ-burst. Because both thermonuclear flashes[14,15] and internal restructuring[16,17] probably originate beneath the outer surface of the neutron star but propagate to the surface in a similar manner, they might be expected to have similar characteristics. However, other features of the source might be used to distinguish between them. In particular, a successful model must be able to function in the presence of a magnetic field, be able to turn on a hard X-ray pulsar, and give recurring bursts with a dynamic range of 100 (refs 3, 18).

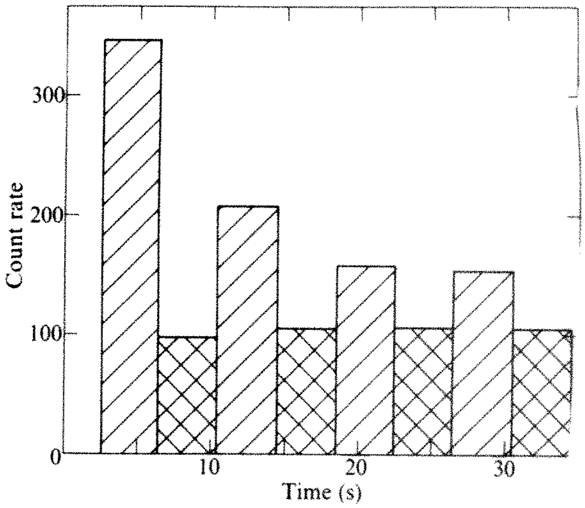


Fig. 2 Average count rate for the X-ray pulses obtained by summing the Venera 11 and 12 observation over 2-s periods (see ref. 3), normalized to the PVO count rate. The cross-hatched regions represent the side of the star associated with the γ-burst and the diagonally-striped regions represent the opposite side.

In the context of internal restructuring, a spin axis different from the magnetic axis (as implied by the pulsations) can apply external torques to the crust of the neutron star which builds up strain even for a slowly rotating neutron star (see review in ref. 19). The resulting starquake would alter the magnetosphere, perhaps providing the short-lived instability driving the X-ray pulsar by accretion. Current models of thermonuclear flashes have reached the quantitative stage for unmagnetized stars but not for magnetized stars. It has been suggested[15,20] that a magnetic field inhibits the flashes and that is why X-ray bursters show no sign of rotational modulation[15]. Because neither internal restructuring models nor magnetized neutron star models have yet reached a stage where quantitative comparisons can be made, the mechanism of the 5 March 1979 γ-burst remains uncertain.

We thank A. Cox, M. Newman, A. Petschek and J. Solem for helpful discussions. This work was funded jointly by NASA and the US Department of Energy.

Received 2 September; accepted 20 October 1980.

1. Klebesadel, R. W., Strong, I. B. & Olson, R. A. *Astrophys. J. Lett.* **182**, L85–88 (197
2. Evans, W. D. *et al. IAU Circ.* No. 3356 (1979).
3. Mazets, E. P. *et al. Nature* **282**, 587–589 (1979).
4. Barat, C. *et al. Astr. Astrophys.* **79**, L24–25 (1979).
5. Cline, T. L. *et al. Astrophys. J. Lett.* **237**, L1–6 (1980).
6. Evans, W. D. *et al. Astrophys. J. Lett.* **237**, L7–10 (1980).
7. Terrell, J., Evans, W. D., Klebesadel, R. W. & Laros, J. G. *et al. Nature* 2 (1980).
8. Evans, W. D. *et al. Science* **205**, 119–121 (1979).
9. Helford, D. J. & Long, S. L. *Nature* **282**, 589–591 (1979).
10. Petschek, A. & Colvin, J. D. *Astrophys. J. Lett.* (submitted).
11. Newman, M. J. & Cox, A. N. *Astrophys. J.* (submitted).
12. Lamb, D. Q. *et al. Nature phys. Sci.* **246**, 52–54 (1973).
13. Lamb, D. Q. & Lamb, F. K. *Ann. NY Acad. Sci.* **302**, 261–299 (1978).
14. Woosley, S. E. & Taam, R. E. *Nature* **263**, 101–103 (1976).
15. Joss, P. C. *Nature* **270**, 310–314 (1977); *Astrophys. J. Lett.* **225**, L1
16. Pacini, F. & Ruderman, M. *Nature* **251**, 399–400 (1974).
17. Fabian, A. C. *et al. Astrophys. Space Sci.* **42**, 77–81 (1976).
18. Mazets, E. P. & Golenetskii, S. V. *Recent Results of the Gamma Burst Studies in the Konus Experiment* (A. F. Ioffe Physico-Technical Institute P 632, 1979).
19. Ruderman, M. A. *Rev. Astr. Astrophys.* **10**, 427–476 (1972).
20. Joss, P. C. & Li, F. K., *Astrophys. J.* **238**, 287–295 (1980).

Is the light cylinder the site of emission in pulsars?

Houshang Ardavan

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

There is no agreement in reports on the oblique-rotator model of pulsars[1-9] as to the site of emission within the pulsar magnetosphere. It is generally assumed, however, that the plasma-electromagnetic fields constituting the magnetosphere obey the quasistatic constraint $\partial f/\partial t = -\omega_0 \partial f/\partial \varphi$, where ω_0 is the angular frequency of rotation of the central neutron star, t the time, and the cylindrical coordinates (r, φ, z) refer to an inertial frame whose z-axis is coincident with the axis of rotation of the star. Accordingly, the pulsar emission is attributed to a source whose distribution has a rigidly rotating pattern; though, of course, none of the charged particles which constitute the source and provide the medium for the propagation of the pattern associated with the source, are themselves constrained to be co-rotating. Here, the spectral distribution of the power emitted by such a rotating pattern is calculated and it is shown that electromagnetic radiation of frequency $\omega \gg \omega_0$, the observed radiation from pulsars, is produced predominantly by the electric currents at the light cylinder $r = c/\omega_0$ of the magnetosphere. This is a model-independent result; it is valid irrespective of the type of emission mechanism and applies to all previously proposed models of the magnetosphere which are quasistatic and in which the radial length scale of the plasma distribution is of the order of c/ω_0.

From the retarded solution of the inhomogeneous Maxwell's equations it follows[10] that the energy radiated per unit frequency and per unit solid angle by any localized distribution of electric current is

$$\frac{\mathrm{d}I(\omega)}{\mathrm{d}\Omega} = \frac{\omega^2}{4\pi^2 c^3}\left|\int \mathrm{d}t \int \mathrm{d}^3x\, \hat{\mathbf{n}} \times [\hat{\mathbf{n}} \times \mathbf{j}(\mathbf{x}, t)] \exp[\mathrm{i}\omega(t - \hat{\mathbf{n}}\cdot\mathbf{x}/c)]\right|^2 \tag{1}$$

where $\mathbf{j}$ is the current density, $\hat{\mathbf{n}}$ is the unit vector directed from the position of the source to the observation point, and the domain of integration extends over all values of the space-time coordinates $\mathbf{x}$ and t. In the inertial coordinate system whose origin is coincident with the position of the neutron star and whose z-axis lies along the rotation axis $\hat{\mathbf{k}}$, an arbitrarily fixed direction of $\hat{\mathbf{n}}$ may be designated by the spherical angles (θ_0, φ_0) where $\theta_0 = \cos^{-1}(\hat{\mathbf{n}}\cdot\hat{\mathbf{k}})$. If in addition the vector $\mathbf{j}$ is decomposed into its cylindrical components (j_r, j_φ, j_z) and $\mathbf{x}$ is expressed in terms of (r, φ, z), then equation (1) assumes the form

$$\frac{\mathrm{d}I(\omega)}{\mathrm{d}\Omega_0} = \frac{\omega^2}{4\pi^2 c^3}|A_\|(\omega, \theta_0, \varphi_0)\hat{\mathbf{e}}_\| + A_\perp(\omega, \theta_0, \varphi_0)\hat{\mathbf{e}}_\perp|^2 \tag{2}$$

$$= \int \mathrm{d}t \int r\,\mathrm{d}r\,\mathrm{d}z\,\mathrm{d}\varphi\, j_{\|,\perp}(r, \varphi, z, t) \times \exp\left\{\mathrm{i}\omega\left[t - \frac{r}{c}\sin\theta_0\cos(\varphi - \varphi_0) - \frac{z}{c}\cos\theta_0\right]\right\} \tag{3}$$

$$j_\| = j_r \sin(\varphi - \varphi_0) + j_\varphi \cos(\varphi - \varphi_0) \tag{4}$$

and the $\theta_0[j_\varphi \sin(\varphi - \varphi_0) - j_r \cos(\varphi - \varphi_0)] + j_z \sin\theta_0$ (5)

$\hat{\mathbf{e}}_\perp = \hat{\mathbf{n}} \times \hat{\mathbf{e}}$ns $\|$ and $\perp$ are defined by $\hat{\mathbf{e}}_\| = (\hat{\mathbf{k}} \times \hat{\mathbf{n}})/|\hat{\mathbf{k}} \times \hat{\mathbf{n}}|$ and

The qua. the cylindric constraint can now be introduced by requiring of φ and t in nponents of the current density to be functions that is by writi ombination $\varphi - \omega_0 t \equiv \hat{\varphi}$ only (see refs 1, 11),

$$j_{r,\varphi,z}(r, \varphi, z, t) = j_{r,}(r, z, \hat{\varphi}) = \int_{-\infty}^{+\infty} j_{r,\varphi,z}(r, z, \hat{\varphi}')\delta(\hat{\varphi}' - \hat{\varphi})\,\mathrm{d}\hat{\varphi}' \tag{6}$$

The second part of equation (6) is no more than a mathematical identity in view of the definition of the Dirac delta function. It provides a useful representation of the quasistatic functions (j_r, j_φ, j_z), however, because by making the integrals over φ and t in equation (3) independent of $\mathbf{j}$, this representation enables us to derive general expressions for $A_\|$ and $A_\perp$ in which the two variables φ and t are everywhere replaced by the single variable $\hat{\varphi}' = \varphi' - \omega_0 t'$. In the case of $A_\|$, for instance, once equation (6) is inserted in equation (3) the integration over t may be immediately performed to obtain

$$\begin{aligned} A_\| = {} & \frac{1}{\omega_0}\exp\left(\mathrm{i}\frac{\omega}{\omega_0}\varphi_0\right)\int r\,\mathrm{d}r\,\mathrm{d}z\,\mathrm{d}\hat{\varphi}' \\ & \times \exp\left[-\mathrm{i}\frac{\omega}{\omega_0}\left(\hat{\varphi}' + \frac{z\omega_0}{c}\cos\theta_0\right)\right]\left\{j_r(r, z, \hat{\varphi}')\int_0^{2\pi}\mathrm{d}\sigma \sin\sigma\right. \\ & \times \exp\left[\mathrm{i}\frac{\omega}{\omega_0}\left(\sigma - \frac{r\omega_0}{c}\sin\theta_0\cos\sigma\right)\right] \\ & + j_\varphi(r, z, \hat{\varphi}')\int_0^{2\pi}\mathrm{d}\sigma\cos\sigma \\ & \left.\times \exp\left[\mathrm{i}\frac{\omega}{\omega_0}\left(\sigma - \frac{r\omega_0}{c}\sin\theta_0\cos\sigma\right)\right]\right\} \end{aligned} \tag{7}$$

where $\sigma \equiv \varphi - \varphi_0$. The integrals over σ can also be carried out explicitly. They are non-zero only when the ratio ω/ω_0 is an integer (ref. 10, p. 96), and are expressible in terms of Bessel functions of integer order m when $\omega = m\omega_0$ (ref. 12, p. 209). Because $\mathbf{j}(r, z, \hat{\varphi})$ is left unchanged whether φ is increased by 2π or equivalently t is decreased by $2\pi/\omega_0$, it is not unexpected that the spectral distribution of the radiation should in the present case be discrete.

The quasistatic versions of $A_\|$ and $A_\perp$ thus obtained, when inserted in equation (2), yield the following expression for the angular distribution of power radiated into the mth harmonic:

$$\frac{\mathrm{d}P_m}{\mathrm{d}\Omega} = \frac{m^2\omega_0^2}{2\pi c^3}|C_\|(m, \theta)\hat{\mathbf{e}}_\| - C_\perp(m, \theta)\hat{\mathbf{e}}_\perp|^2 \tag{8}$$

where

$$C_\| = \int r\,\mathrm{d}r\,\mathrm{d}z\,\mathrm{d}\hat{\varphi}\exp\left[-\mathrm{i}m\left(\hat{\varphi} + \frac{z\omega_0}{c}\cos\theta\right)\right] \times \left[j_\varphi\frac{dJ_m(m\xi)}{d(m\xi)} + \frac{\mathrm{i}}{\xi}j_r J_m(m\xi)\right] \tag{9}$$

$$C_\perp = \int r\,\mathrm{d}r\,\mathrm{d}z\,\mathrm{d}\hat{\varphi}\exp\left[-\mathrm{i}m\left(\hat{\varphi} + \frac{z\omega_0}{c}\cos\theta\right)\right] \times \left[j_r\cos\theta\frac{dJ_m(m\xi)}{d(m\xi)} - \mathrm{i}\left(\frac{\cos\theta}{\xi}j_\varphi - \sin\theta j_z\right)J_m(m\xi)\right] \tag{10}$$

and

$$\xi \equiv \frac{r\omega_0}{c}\sin\theta$$

Here, $\hat{\varphi}'$ and θ_0 have been written $\hat{\varphi}$ and θ for brevity, and $\mathrm{d}I/\mathrm{d}\Omega$ has been replaced by $(2\pi/\omega_0^2)(\mathrm{d}P_m/\mathrm{d}\Omega)$ to convert energy per unit frequency to power per harmonic. Note that j_r, j_φ and j_z are functions of $(r, z, \hat{\varphi})$, and that the integrals extend over all values of r, z and $\hat{\varphi}$ for which the current density is non-zero. It is easy to see that when specialized to the case of a point source for which $j_r = j_z = 0$, equation (8) is the same as the corresponding expression for synchrotron radiation[12].

In the range of frequencies $m \gg 1$, asymptotic values of the Bessel function appearing in equation (8) are given by[13]

$$J_m(m\,\mathrm{sech}\,\alpha) \sim (2\pi m\tanh\alpha)^{-1/2}\exp[m(\tanh\alpha - \alpha)] \tag{11}$$

or

$$J_m(m\sec\beta) \sim (\tfrac{1}{2}\pi m\tan\beta)^{-1/2}\cos[m(\tan\beta - \beta) - \tfrac{1}{4}\pi] \tag{12}$$

0028-0836/81/020044—02$0

depending on whether ξ is greater or less than one. When the current density has limited total fluctuation in the radial direction, therefore, the rapid oscillation of the Bessel function in the interval $\xi > 1$ renders the magnitude of the power originating from $(r\omega_0/c)\sin\theta > 1$ negligible. According to the principle of stationary phase, the main contributions to the values of the integrals in equation (8) for $\xi > 1$ come from the neighbourhood of the surface where

$$\frac{\mathrm{d}}{\mathrm{d}\beta}[m(\tan\beta-\beta)-\tfrac{1}{4}\pi]=m\left[\left(\frac{r\omega_0}{c}\sin\theta\right)^2-1\right]=0 \qquad (13)$$

A similar conclusion in fact follows from the behaviour of the Bessel function in $\xi < 1$. Equation (11) in conjunction with the asymptotic value of the Bessel function at $\xi = 1$ (ref. 13) shows that the ratio

$$\frac{J_m(m\operatorname{sech}\alpha)}{J_m(m)} \sim \frac{3^{2/3}\Gamma(\frac{2}{3})}{2^{1/3}(2\pi)^{1/2}}\,\frac{\exp[m(\tanh\alpha-\alpha)]}{m^{1/6}\tanh\alpha} \qquad (14)$$

has a sharp peak in the neighbourhood of $(r\omega_0/c)\sin\theta = 1$. In this case, the order of magnitude $m^{-1/6}\exp(-\alpha m)$ of this ratio for $\alpha \gg 1$, that is for $r\omega_0/c \ll 1$, is small for the same mathematical reason that the power radiated by a particle in the non-relativistic as compared with that in the extreme relativistic regime is small. This mathematical similarity is, of course, not fortuitous (see ref. 11).

Equation (8) also shows that for a current density whose length scale of variation along z is c/ω_0, the main contribution to the values of $C_\parallel$ and $C_\perp$ is received from directions where $m\cos\theta \simeq 1$, because the integrals over z simply represent the Fourier amplitudes of the current density corresponding to the wave number $m\omega_0\cos\theta/c$. This beaming of the high-frequency radiation into $\theta \simeq \pi/2$, in turn implies that ξ attains the value unity in the neighbourhood of the light cylinder. Therefore I suggest that the higher the frequency, the closer the site of emission to the light cylinder. Because the quasistatic condition is associated with a rotating pattern for the source (see ref. 11 for a detailed discussion), it is not too surprising that the emission from inside the light cylinder should be much smaller than that from the vicinity of the light cylinder; the more unexpected outcome of the present analysis is that the emission from beyond the light cylinder is also relatively small.

Because it follows from a formalism in which the electric current is left unspecified, the present result is clearly independent of both the details of the mechanism by which the radiation is emitted and of the effects of the plasma through which it is propagated. The structure of the magnetosphere, moreover, has a bearing on this result only in that the plasma distribution is required to have a length scale of the order of c/ω_0. The validity of the quasistatic condition is, of course, implicitly dependent on the existence of an agent—such as a large scale magnetic field—which affects the coupling between the neutron star and the magnetosphere. However, imposing the quasistatic constraint can also be justified on independent observational grounds[11].

The present analysis clearly establishes that there is an inconsistency in those models which adopt the quasistatic constraint and at the same time identify the surface of the neutron star as the site of emission. According to equation (14), the power radiated into the mth harmonic $\omega = m\omega_0$ receives a contribution from the electric currents in the vicinity of the light cylinder which is by a factor of the order of $m^{1/3}\exp(2\alpha m)$ greater than that received from the vicinity of the neutron star. Thus, even for radio waves with $m \simeq 10^8$, the observed emission originates almost exclusively from the light cylinder when the electric currents are distributed throughout the magnetosphere.

The present discussion has centred on only one particular implication of equation (8), but in fact this equation, which incorporates the coherent character of the radiation automatically, also constitutes the departure point for a study of the polarization and the spectrum of the pulsar emission.

Received 11 August; accepted 6 November 1980.

1. Mestel, L. *Nature phys. Sci.* **233,** 149–152 (1971).
2. Cohen, J. M. & Rosenblum, A. *Astrophys. Space Sci.* **16,** 130–136 (1972).
3. Roberts, D. H. & Sturrock, P. A. *Astrophys. J.* **181,** 161–180 (1973).
4. Endean, V. G. *Astrophys. J.* **187,** 359–360 (1974).
5. Henriksen, R. N. & Norton, J. A. *Astrophys. J.* **201,** 719–728 (1975).
6. Mestrel, L., Wright, G. A. E. & Westfold, K. C. *Mon. Not. R. astr. Soc.* **175,** 257–278 (1976).
7. Cheng, A. F. & Ruderman, M. A. *Astrophys. J.* **212,** 800–806 (1977).
8. Buckely, R. *Mon. Not. R. astr. Soc.* **183,** 771–778 (1978).
9. Arons, J. & Scharlemann, E. T. *Astrophys. J.* **231,** 854–879 (1979).
10. Jackson, J. D. *Classical Electrodynamics* (Wiley, New York, 1962).
11. Ardavan, H. *Astrophys. J.* (submitted).
12. Landau, L. D. & Lifshitz, E. M. *The Classical Theory of Fields* (Pergamon, Oxford, 1962).
13. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions* (Dover, New York, 1965).

Photovoltaic properties of merocyanine solid-state photocells

G. A. Chamberlain & P. J. Cooney

Shell Research Ltd, Thornton Research Centre, PO Box 1,
Chester CH1 3SH, UK

S. Dennison

Department of Physical Chemistry, Lensfield Road, University of
Cambridge, Cambridge CB2 1EP, UK

Considerable progress has been made in improving the sunlight conversion efficiency of solid-state organic photovoltaic cells. Ghosh *et al.*[1] report a 0.7% efficiency for a merocyanine dye Schottky barrier cell. We expect that improvements in energy conversion efficiency may follow if merocyanines of low ionization potential are used in photovoltaic cells, in which the dye layer contains a strongly electronegative dopant such as iodine. To test these ideas we report here the photovoltaic quantum yields and sunlight conversion efficiences of a series of merocyanines in which the molecular properties have been varied systematically.

Ghosh *et al.*[1] used an evaporated thin film of the dye, sandwiched between thin Al/Al_2O_3 and silver layers and illuminated through the Al contact for maximum response. They proposed that charge carriers are generated by electron transfer from an exciton to unoccupied metal orbitals or oxide surface states at the aluminium/dye interface. The magnitude of the photocurrent at a given light intensity is then governed by the field dependence of geminate recombination between the transferred electron on the aluminium and its positive image charge in the dye layer. We have observed, however, that the photovoltaic response of this type of cell is extremely sensitive to the presence of added gases, for example, ambient air and particularly iodine vapour. By preparing and testing the cell in a vacuum of 10^{-6} torr we have found that the magnitude of the photovoltaic response increases by several powers of 10 when the dye is doped with iodine vapour; the dark conductivity of the dye, however, is unchanged. It seems likely, therefore, that charge carriers may be photogenerated by a charge transfer process from excited dye to iodine, possibly through an exciplex precursor. The undesirable back electron transfer reaction must then compete with charge carrier separation by hopping, which occurs in the built-in field in the region of the aluminium/dye interface. The extent of charge transfer from a donor (the dye) to an acceptor (iodine) in the ground and excited states of the complex increases as the ionization potential of the donor decreases[2].

In general terms[3], a merocyanine dye consists of an electron-donating group (the donor) linked by a conjugated polymethine bridge to an electron-accepting group (the acceptor). The intramolecular charge transfer that occurs in the ground and excited states depends on the relative strengths of the donor and acceptor groups. In addition, the stability and lifetime of the

0028–0836/81/020045—03$01.00

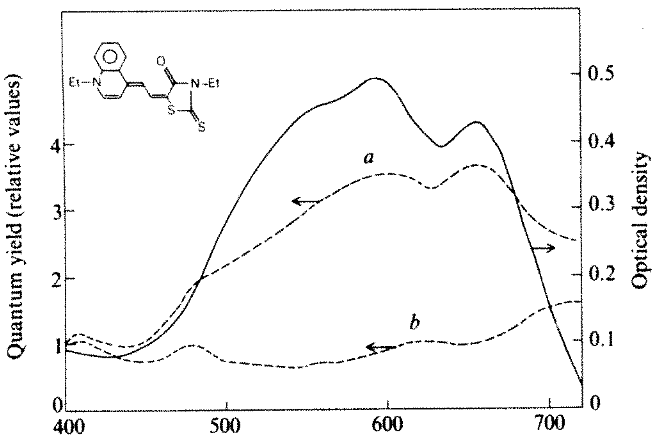


Fig. 1 Absorption spectrum and photocurrent action spectra for dye *E*-Ac, ~35 nm thick. *a*, Illumination through Al; *b*, illumination through Au.

cationic form of the dye that is produced after intermolecular charge transfer in photovoltaic cells depends on the ease with which the positive charge can be delocalized. This is determined to a large extent by the ionization potential of the dye, which in turn depends on the donating ability of the donor fragment. Therefore, to assess the effect of donor strength and bridge length on photovoltaic behaviour we have screened the responses of several merocyanines containing the same acceptor group (ethyl rhodanine). Some interesting molecular structure/photovoltaic activity relationships have emerged.

Solid-state sandwich cells were prepared by the sequential vacuum evaporation and deposition of thin films of aluminium, aluminium oxide, merocyanine dye and gold onto a quartz or Pyrex substrate supplied with two indium contacts. Active areas of about 5 cm^2 were readily obtained. The cells were then doped with iodine by vapour diffusion. An area of 0.064 cm^2 was illuminated with a single-strand Crofon fibre optic. Dark and photoelectric responses were monitored throughout the preparation of the cell and the apparatus was arranged such that the vacuum was maintained at about 10^{-6} torr. The thickness of each layer was controlled during evaporation by means of a Kronos quartz crystal thickness monitor. The metals used were at least 99.99% pure, and the purity of the dyes (~99%), prepared at Thornton Research Centre, was checked after double recrystallization by elemental analysis, IR spectroscopy and mass spectroscopy. Redox potentials were obtained from cyclic voltammograms of CH_2Cl_2 solutions of the dyes using Pt electrodes and tetrabutyl ammonium perchlorate as supporting electrolyte against a Ag/AgCl reference electrode.

The merocyanines and data obtained are shown in Table 1. For comparison, in the solid-state cells the thickness of each dye was maintained at about 35 nm, and quantum yields (ϕ_ε) based on photons absorbed by the dye (that is, correcting for transmittance of the aluminium layer) were measured at the peak of the absorption spectrum and in a range of light intensity where the photocurrent varied linearly. For donors A–E, ϕ_ε followed the ease of oxidation of the dye, which correlates well with the expected trends from charge transfer behaviour, resonance theory and the Brooker deviation[4], which give a qualitative measure of the relative donating ability of the donor residue.

Increasing the bridge length reduces the ionization potential and extends the chromophore; compare results for donors E, F and G. Unfortunately, merocyanine G-Ac became unstable during sublimation, and a thermal rearrangement to a non-coloured form, possibly a spiropyran[5], occurred. No photovoltaic effect was observed because the compound did not absorb in the visible portion of the spectrum. The original compound, however, could be regenerated by dissolving the thin film in ethanol, suggesting that a reversible tautomerization had occurred. Merocyanine F-Ac was tested in the hope that removal of the bridge between donor and acceptor groups would reduce the likelihood of tautomerization. However, no photovoltaic effect was observed, even though there was no evidence of oxetane formation. This result may be explained in terms of the excited-state lifetime of the dye. The steric hindrance between donor and acceptor groups leads to loss of coplanarity. Tredwell and Keary[6] have shown that twisting of heterocyclic rings about the polymethine bridge to relieve steric strain enhances the rate of internal conversion to the ground state, because excited state energy can then be efficiently dissipated by intramolecular vibration and rotation. We estimate from their results that the excited state lifetime $(\Sigma k)^{-1}$ for dye F-Ac is of the order of 1 ps. The charge generation process

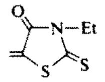
$$(\text{dye})^* + \text{dopant} \xrightarrow{k_d} (\text{dye})^+ + (\text{dopant})^-$$

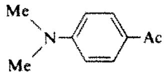
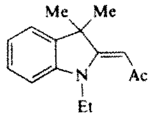
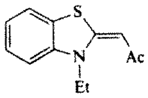
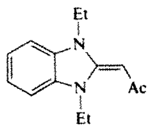
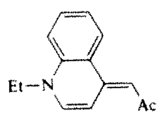
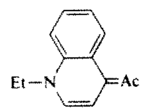
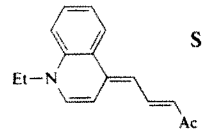
has a first-order rate constant[7] of about $10^7\ \text{s}^{-1}$ for a dopant concentration of 10^{19} molecules cm^{-3}. Thus, the quantum efficiency ϕ_d for excited state dissociation to charge carriers given by:

$$\phi_d = \frac{k_d[\text{dopant}]}{k_d[\text{dopant}] + \Sigma k}$$

has a value of about 0.001%, which is near the limit of detectability in our system. When combined with other energy dis-

Table 1 Oxidation potentials, ionization energies[7] and photovoltaic quantum yields at λ_{max} for merocyanine dyes derived from ethyl rhodanine, Ac


Ethyl rhodanine (Ac)

Merocyanine	Relative strength of donor	Oxidation potential (V/NHE)	Ionation potential (eV)	λ_{max} (nm)	ϕ_ε %
A 	Weak	1.29	5.8	420	<0.001
B 	Weak	0.92*	5.4	520	0.08
C 	Medium	0.98*	5.5	550	10
D 	Strong	0.73*	5.2	540	14
E 	Strong	0.77*	5.3	660	12
F 	Strong	0.99	5.5	550	0
G 	Strong	0.57*	5.1	770	Decomposed

* Irreversible, peak values recorded.

Table 2 Sunlight conversion efficiencies η_s and dependence of photocurrent I_{sc} on light intensity for Al/Al_2O_3/merocyanine, I_2/Au photovoltaic cells

Merocyanine	I_o (mW cm^{-2})	% Transmittance of Al	η_s %	I_{sc}, I_o dependence
A	180*	50	$\sim 10^{-6}$	$I_{sc} \propto I_o^{0.5}$
B	63	65	2×10^{-5}	$I_{sc} \propto I_o^{10.5}$
C	52	30	0.013	$I_{sc} \propto I_o$
D	56	10	0.007	$I_{sc} \propto I_o$
E	45	27	0.18	$I_{sc} \propto I_o$
F	52	25	No response	—

I_o is intensity of simulated AM2 sunlight incident on the Al electrode.
* Tungsten-halogen lamp.

sipation processes, such as carrier recombination, this estimate accounts for the absence of detectable photocurrent.

The photocurrent action spectra for dye E-Ac are shown in Fig. 1. Good correspondence between action spectrum and absorption spectrum is observed for illumination through the aluminium electrode. However, an inverted behaviour is apparent when the cell is illuminated through the gold side. These results are consistent with the formation of a rectifying junction at the aluminium/dye contact. In the case of gold illumination, the dye itself acts as a filter of the strongly absorbed light, and weakly absorbed photons at the sides of the absorption peaks penetrate to the rectifying junction and make a major contribution to the photocurrent.

The dyes were tested for their photovoltaic power conversion efficiency, η_s, in simulated AM2 sunlight from a Thorn CSI lamp. η_s was calculated from the short-circuit current (I_{sc}), open-circuit voltage (V_{oc}) and fill factor (FF) at known incident light intensity (I_o) using the equation:

$$\eta_s = I_{sc} V_{oc} \mathrm{FF} / I_o$$

and the results are shown in Table 2. Efficiencies were dependent not only on donor type but also on the relationship between I_{sc} and I_o. The more efficient dyes, C-Ac, D-Ac and E-Ac, had linear photocurrent dependences up to and beyond the light intensities used. Taking the response for the most efficient dye, E-Ac, for a film of about 25 nm thickness, we have estimated the power conversion efficiency based on light absorbed, η_s', defined as:

$$\eta_s' = \frac{I_{sc}}{e\int_{E_g}^{E_m} A(E)N(E)\mathrm{d}E} \cdot \frac{eV_{oc}}{E_g} \cdot \mathrm{FF} \times 100\%$$

$$= [\text{average quantum yield}][\text{voltage factor}][\text{fill factor}] \times 100\%$$

where e is the electronic charge, E_m is the high-energy limit of the solar spectrum (taken as 3.1 eV), E_g is the absorption threshold for the dye (taken as 1.7 eV), $A(E)$ is a factor which accounts for the dye absorbance and $N(E)$ is the flux density of photons incident on the dye film per wavelength interval. From estimates of the average quantum yield (0.13), voltage factor (0.42) and measured fill factor (0.39) we calculate $\eta_s' = 2.1\%$.

Further exploitation of the relationship between molecular structure and photovoltaic behaviour should lead to further improvements in the conversion efficiency.

We thank G. Ellis for preparing pure samples of the merocyanine dyes.

Received 16 July; accepted 7 November 1980.

1. Ghosh, A. K. & Feng, T. *J. appl. Phys.* **49**, 5982–5989 (1978).
2. Birks, J. B. in *Photophysics of Aromatic Molecules* Ch. 9 (Wiley, London, 1970).
3. Griffiths, J. in *Colour and Constitution of Organic Molecules* Ch. 6. (Academic, London, 1976).
4. Brooker, L. G. S. *et al. J. Am. chem. Soc.* **73**, 5332–5350 (1951).
5. Bertelson, R. C. in *Techniques of Chemistry* Vol. 3, Ch. 3 (ed. Brown, G. H.) (Wiley, London, 1971).
6. Tredwell, C. J. & Keary, C. M. *Chem. Phys.* **43**, 307–316 (1979).
7. Chamberlain, G. A. & Malpas, R. E. *JCS Faraday Discuss.* **70** (in the press).

0028–0836/81/020047—03$01.00

Chemical changes during dyke metamorphism in high-grade basement terrains

Barry L. Weaver & John Tarney

Department of Geology, The University of Leicester,
Leicester LE1 7RH, UK

Trace element geochemistry is an important tool in deciphering the petrogenesis of igneous, and particularly basaltic, igneous rocks, and in assigning a possible tectonic setting to these rocks[1,2]. Studies of the trace element geochemistry of metamorphosed igneous rocks often implicitly assume element immobility during metamorphism even though recent studies[3–7] have demonstrated the sometimes quite extreme mobility of certain trace elements, and especially of the petrogenetically important rare-earth elements (REE), during low-grade (zeolite–greenschist facies) metamorphism. The mobility of trace elements at higher grades of metamorphism is poorly known, although there is some evidence[8] for the mobility of large ion lithophile (LIL—Rb, K, Th, Ba, Sr) elements, but not the REE, during amphibolite facies metamorphism. The REE also seem to be stable through the amphibolite to granulite facies transition[9,10]. We have studied a suite of ultramafic to mafic dykes (the Scourie dykes) cutting regrogressed Lewisian granulites in the Assynt region, north-west Scotland[11] and conclude that the mobility of trace elements is not controlled solely by the degree or grade of metamorphism, but more by the availability of a fluid phase for element transfer.

The intrusion of the Scourie dolerite dykes has been dated at 2,390 Myr (ref. 12). In the Assynt region, the dyke suite may have been intruded over a time span as great as 200 Myr (ref. 13), intrusion of the early dolerites being followed by intrusion of ultramafic dykes and later, rather minor, dolerites[11]. There is evidence for the injection and metamorphism of these dykes at depth in the crust (under pressures of 5–7 kbar) into hot (~500 °C) country rock gneisses, during the waning stages of the Inverian retrogression of Scourian granulites[11,14–16]. The dykes are variably metamorphosed, with the earlier dykes being in general more heavily metamorphosed[11,14]. The metamorphism of dolerite dykes from the Scourie area has been described by Sutton and Watson[17], who recognized six progressive stages of metamorphism. Metamorphic conditions in Assynt (of the lower amphibolite facies) were somewhat lower than at Scourie[11]. The metamorphism of Assynt ultramafic dykes has been detailed elsewhere[11].

The present study of the chemical changes associated with the metamorphism of Scourie dykes arises from the analysis of a large number of dykes from Assynt for a wide range of elements[18]. Amongst the trace elements, the incompatible elements are of most use in petrogenetic modelling, and it is important to assess the behaviour of these elements during metamorphism. As both fresh and variably metamorphosed dykes are present in Assynt, the behaviour of incompatible elements on metamorphism can be gauged by the normalization of the incompatible element abundances in an amphibolized dyke to those in a cogenetic fresh dyke. Ignoring possible effects due to phenocryst mineralogy, the resulting incompatible element pattern should be linear on such a plot if mobilization of trace elements has not occurred. The anomalous behaviour of an element on such a plot will, however, reflect, if phenocryst effects can be discounted, mobility of that element during metamorphism. It has been demonstrated[18] that during fractionation of the cogenetic fresh Scourie dolerites, the levels of incompatible elements increase steadily, with little change in incompatible element ratios, except for Sr, which has an approximately constant abundance due to the plagioclase fractionation which these dykes seem to have undergone. Bearing this in mind, the metamorphic effects on the chemistry of Scourie dolerites is illustrated in Fig. 1*a*, where the incompatible element abundances (arranged in the order given in ref. 19) in both a fresh and an amphibolized dolerite dyke have been normalized relative to a fresh dolerite. The fresh dolerite (sample 916) on this plot (Fig. 1*a*) displays an overall flat pattern, but with a negative Sr anomaly and a positive Ti anomaly which can be attributed to crystal fractionation effects[18]. On the other hand, the amphibolized dolerite (sample 689, Fig. 1*a*), in which pyroxene has been completely replaced by hornblende–quartz aggregates, the feldspar is altered and recrystallized, and the original texture of the rock has disappeared, displays significant relative depletion in Sr, Ba, K and Rb. The Sr anomaly can again be attributed to crystal fractionation[18], but the depletion in Ba, K and Rb is a direct result of amphibolite facies metamorphism. To assess at what stage in the metamorphic process appreciable K and Rb loss occurs, the analysed dolerites have been divided into three groups:

(1) those dykes which are fresh, with little or no replacement of pyroxene by secondary amphibole (nine samples);
(2) those dykes in which there is complete replacement of pyroxene, but with textural preservation (four samples);
(3) those dykes in which recrystallization of feldspar and the hornblende–quartz aggregates has occurred with destruction of the original texture (eight samples).

The average K_2O and Rb contents for the three groups are 0.82% and 18 p.p.m., 0.92% and 18 p.p.m., and 0.66% and 8 p.p.m. respectively, yielding average K/Rb ratios for each group, with 1σ standard deviations, of 392 ± 63, 422 ± 115 and 759 ± 356. Clearly then, substantial loss of K and Rb (and other LIL elements), and the consequent development of high K/Rb ratios, occurs only on recrystallization of the dykes.

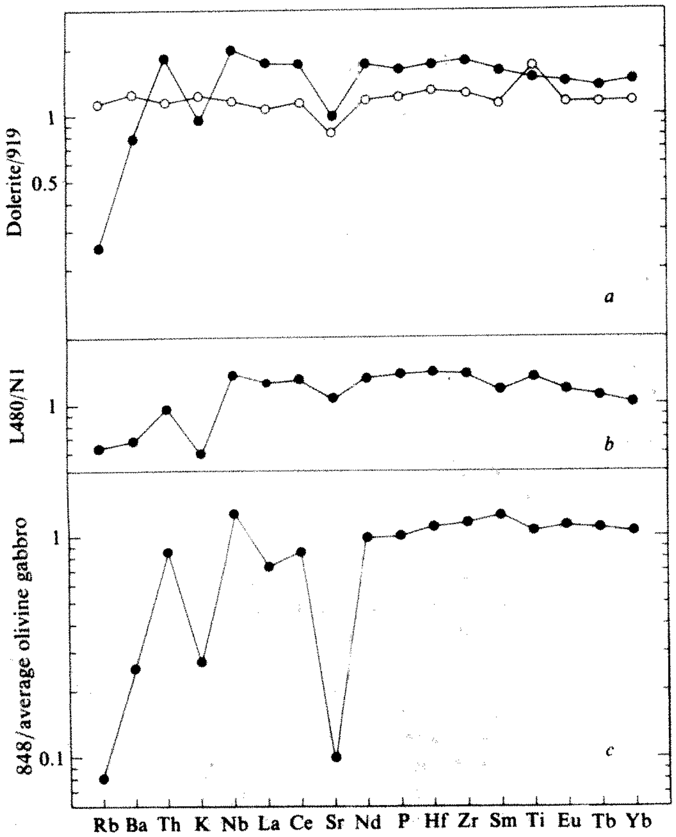


Fig. 1 Plot of incompatible element abundances in: *a*, fresh dolerite (○, 916) and an amphibolized dolerite (●, 689) normalized to a fresh dolerite dyke (919); *b*, an altered 'norite' dyke (L480) normalized to a fresh 'norite' dyke (N1); and *c*, an altered olivine gabbro dyke (848) normalized to the average of four fresh olivine gabbros.

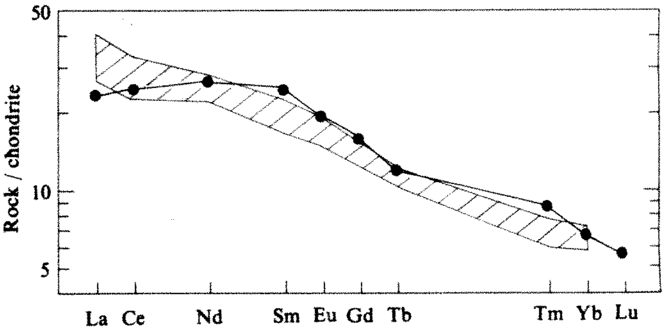


Fig. 2 Chondrite normalized plot of REE abundances in an altered olivine gabbro (●, 848) compared with the field for four fresh olivine gabbro dykes (from ref. 18). Chondrite normalizing values used are those of Nakamura[23].

Figure 1*b*, *c* investigates the metamorphism of ultramafic Scourie dykes. In Fig. 1*b*, sample L480 (the Loch Crocach 'norite' dyke[11]), which is heavily amphibolized throughout its length, has been normalized to a cogenetic fresh norite dyke from Scourie (dyke N1 of ref. 15). Here the LIL elements K, Rb, Ba and, less clearly, Sr seem to have been mobilized (Fig. 1*b*) during metamorphism of this ultramafic dyke. However, the depletion of these elements is not particularly severe, perhaps because sample L480 shows a considerable degree of textural preservation, with fresh primary feldspar and pseudomorphs after pyroxene still obvious. On the other hand, Fig. 1*c* illustrates the chemical effects of metamorphism of the olivine gabbro dykes from Assynt[11]. Sample 848 is, on the basis of most major and trace elements, clearly related to the olivine gabbro dykes[18], although the original olivine–augite–orthopyroxene–plagioclase–amphibole–mica assemblage[11] has been totally replaced by an amphibole–chlorite–ore mineralogy, with no textural preservation. This extreme metamorphism has caused severe depletion in Sr, Ba, K and Rb in the amphibolized olivine gabbro, while Th, La and Ce are less strongly depleted. The depletion in the light REE is perhaps better illustrated in Fig. 2, where the REE abundances in 848 are compared with the range in REE patterns for four fresh olivine gabbro dykes[18].

Apparently then, amphibolite facies metamorphism of Scourie dykes leads to appreciable mobilization of certain trace elements on recrystallization of the dyke. In this case, depletion in the LIL elements and the LREE is demonstrable, with mobility decreasing in the approximate order: $\mathrm{Rb} > \mathrm{Ba} \simeq \mathrm{K} \simeq \mathrm{Sr} \gg \mathrm{Th} \simeq \mathrm{La} \simeq \mathrm{Ce}$, although depletion in Th, La and Ce only becomes significant on extreme amphibolization. It would also seem reasonable that the mobility of Sr will depend on the stability of feldspar: in those rocks which retain feldspar on amphibolization (the dolerite dykes and the Loch Crocach norite), Sr is not particularly mobile, while in the amphibolized olivine gabbro, which contains no relict feldspar, Sr has been severely depleted. On the other hand, the elements Nb, P, Hf, Zr, Ti and the middle- and heavy-REE are immobile for all degrees of metamorphism observed in the Scourie dykes.

The critical factor seems to be not the grade of metamorphism, but the availability of a fluid agency for element transfer. During the amphibolite-facies retrogression of the enclosing granulite-facies gneisses (which largely preceded, but overlapped, the period of dyke injection), significant chemical changes as the gneiss compositions readjusted to their new equilibrium mineral assemblages have been recorded[20]. In this case, very considerable volumes of hydrous fluids were available to assist with element redistribution. The dykes, being intruded towards the end of this period of retrogression, were metamorphosed to a degree dependent on the availability of hydrous fluids. Thus they are metamorphosed progressively from their margins to their centres[11]. The dyke margins, which are often sheared, provided the means of access of these hydrous fluids from below[20]. High metamorphic temperatures alone were not sufficient to cause recrystallization, and only when the dykes had been thoroughly metamorphosed, implying a large fluid flux, are there significant chemical changes. Conversely, low-grade metamorphism of ocean-floor basalts is frequently associated with quite large chemical changes[21], even affecting rare-earth elements[22]; however, this is frequently related to the action of hydrothermal convection cells at mid-ocean ridges in which the fluid flux is of some magnitude[21], and high water–rock ratios are maintained for a considerable time[22].

Metamorphism alone is not sufficient to initiate chemical changes; the fluid flux is a much more important factor, but one which is rather more difficult to assess. The metamorphic grade does, however, control the mineral assemblages, which determines which elements may be mobile.

B.L.W. acknowledges the receipt of an NERC research studentship. The analytical data used in this study were obtained at the University of Birmingham (X-ray fluorescence data) and the Open University (NAA data), and the use of these facilities is gratefully acknowledged.

Received 11 September; accepted 30 October 1980.

1. Pearce, J. A. & Cann, J. R. *Earth planet. Sci. Lett.* **19,** 290 (1973).
2. Winchester, J. A. & Floyd, P. A. *Earth planet. Sci. Lett.* **28,** 459 (1976).
3. Herrmann, A. G., Potts, M. J. & Knake, D. *Contr. Miner. Petrol.* **44** (1974).
4. Humphris, S. E., Morrison, M. A. & Thompson, R. N. *Chem. Geol.* **23,** 125 (1978).
5. Ludden, J. N. & Thompson, G. *Earth planet. Sci. Lett.* **43,** 85 (1979).
6. Hellman, P. L., Smith, R. E. & Henderson, P. *Contr. Miner. Petrol.* **71,** 23 (1979).
7. Nesbitt, H. W. *Nature* **279,** 206 (1979).
8. Muecke, G. K., Pride, C. & Sarkar, P. *Phys. Chem. Earth* **11,** 449 (1979).
9. Green, T. H., Brunfelt, A. O. & Heier, K. S. *Geochim. cosmochim. Acta* **36,** 241 (1972).
10. Compton, P. *Contr. Miner. Petrol.* **66,** 283 (1978).
11. Tarney, J. in *The Early Precambrian of Scotland and Related Rocks of Greenland* (eds Park, R. G. & Tarney, J.) 105 (University of Keele, 1973).
12. Chapman, H. J. *Nature* **277,** 642 (1979).
13. Evans, C. R. & Tarney, J. *Nature* **204,** 638 (1964).
14. Tarney, J. *Nature* **199,** 672 (1963).
15. O'Hara, M. J. *Miner. Mag.* **32,** 848 (1961).
16. O'Hara, M. J. *J. geol. Soc. Lond.* **134,** 185 (1977).
17. Sutton, J. & Watson, J. V. *Geol. Mag.* **88,** 25 (1951).
18. Weaver, B. L. & Tarney, J. (in preparation).
19. Wood, D. A., Joron, J.-L., Treuil, M., Norry, M. J. & Tarney, J. *Contr. Miner. Petrol.* **70,** 319 (1979).
20. Beach, A. & Tarney, J. *Precamb. Res.* **7,** 325 (1978).
21. Donnelly, T. W., Thompson, G. & Robinson, P. T. *Am. geophys. Un. Maurice Ewing Ser.* **2,** 369 (1979).
22. Menzies, M., Seyfried, W. & Blanchard, D. *Nature* **282,** 398 (1979).
23. Nakamura, N. *Geochim. cosmochim. Acta* **38,** 757 (1974).

An unusual bed of giant pumice in Mexico

B. J. Clough, J. V. Wright*† & G. P. L. Walker*

Department of Geology, Imperial College, London SW7 2BP, UK

We report here that a remarkable bed of giant-sized pumice blocks occurs within an old caldera of the major late Quaternary rhyolitic volcano of Sierra La Primavera, Mexico (Fig. 1*a*). The bed lies within a 100-m thick sedimentary sequence distributed over a near-circular area 10 km across. The margin of this circular area closely parallels an incomplete ring of prophyritic comendite lava domes (Fig. 1*b*), some of which may be shallow intrusions emplaced into the sediments. The vents of these domes are believed to lie on a ring fault (Fig. 1*b*). The sedimentary sequence, now gently updomed (Fig. 1*c*), is believed to have accumulated in a caldera lake, the shoreline of which lay up to 1 km outside the ring fault.

The lake sediments are floored by an ignimbrite (Rio Caliente Ignimbrite) produced by the biggest explosive eruption of the volcano[1,2]; many of the deep gullies within the badland terrain, which characterize the area of lake deposits, cut down into the floor. The lake deposits have a maximum exposed thickness of ~100 m, and rest on top of the ignimbrite which was eroded before their accumulation. Locally, the non-welded top of the ignimbrite (at least 15 m) was stripped off and box canyons were incised into the incipiently welded zone. Some of these canyons are infilled with cross-stratified and pumice-rich fluvatile deposits.

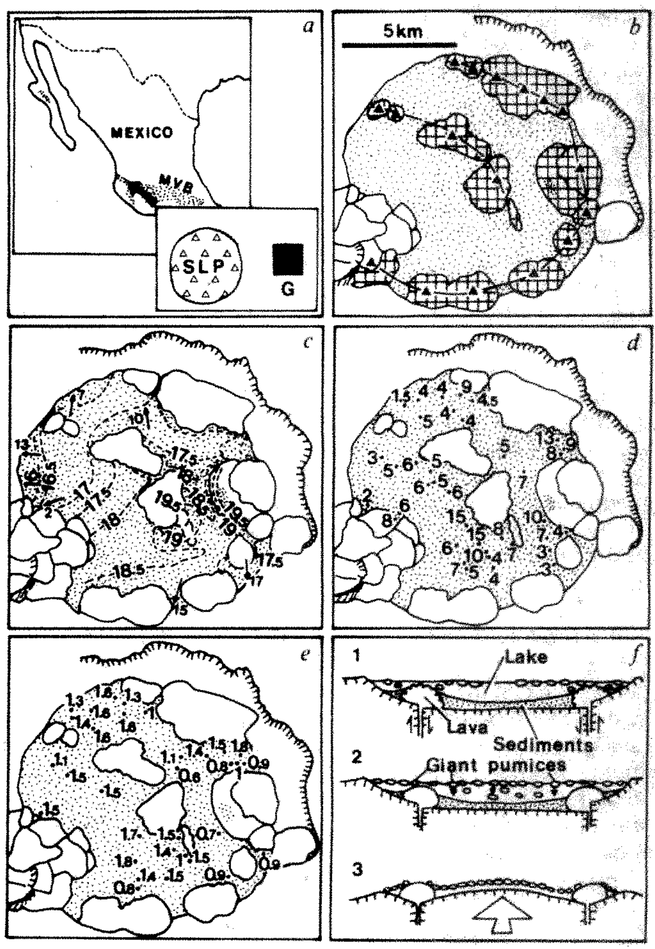


Fig. 1 The giant pumice block bed and lake sediments, and their relationships to the young rhyolite lavas of La Primavera volcano. *a*, Index map of study area: MVB is the Mexican Volcanic Belt which is a locus of Quaternary volcanism in Mexico; G is Guadalajara; SLP is the Sierra La Primavera. *b*, Distribution of the lake sediments and rhyolite lavas: stipple is lake sediments; cross-hatch is porphyritic comendite lavas lying on an 'inner-ring' with a transverse zone; triangles represent inferred vent positions; unornamented lavas are aphyric comendite lavas lying on an 'outer-ring' which is not fully shown or discussed here; hatched line is a low escarpment, possibly an embryo caldera rim. *c*, Altitude contours of the top of the giant pumice block bed in hundreds of metres above sea level. Heights were determined in the field using an altimeter. Dip of lake sediments is also shown (arrows). *d*, Thickness of giant pumice block bed in metres. *e*, Median diameter of giant pumice blocks in metres; 65–20 blocks were measured at each outcrop. *f*, Three stages in the envisaged formation of the giant pumice block bed.

The lake deposits are planar-laminated white vitric ashes and pumice layers. Some ash beds contain diatoms; discrete beds of diatomite also occur. The sediments commonly have current-ripple laminae and load casts indicative of deposition under water. Diatom assemblages indicate that the water was shallow (G. H. Evans, personal communication) and less saline than present-day lakes in this area (J. P. Bradbury, personal communication).

The most distinctive unit in the lake-sediment succession is the giant pumice block bed (Fig. 2*a*), which is 1.5–15 m thick and can be traced over practically the entire former lake area (Fig. 1*d*). It has an extreme bimodal size distribution, the coarse mode consisting of giant blocks of light grey porphyritic pumice averaging 1.3 m in size (Fig. 1*e*), commonly grain supported, with the finer mode consisting of fine white ash occurring

---

* Present addresses: Department of Geology, University of Puerto Rico, Mayaguez, Puerto Rico 00708 (J.V.W.) and Department of Geology, University of Auckland, New Zealand (G.P.L.W.).
† To whom correspondence should be addressed.

0028–0836/81/020049—02$01.00

between the large blocks. The giant pumice blocks (Fig. 2) range in diameter from 0.2 to generally less than 6 m and are thus quite well sorted for their size; the maximum diameter measured of any block was 8.5 m. Locally, the giant pumice block bed is underlain by a thick accumulation of smaller-sized pumice. Near the northern edge of the lake area, the bed occurs as two or three beds separated by stratified reworked ash and pumice. The giant pumice blocks in the upper two beds are more matrix supported than is usual for this deposit, and may be reworked.

The ash matrix in the giant pumice bed is laminated, and the laminations are disturbed and contorted around the pumice (Figs 2*a*, 3); this is interpreted as gentle settling of the pumice blocks into waterlogged ash. However, the laminations are also disturbed above the pumice blocks, which is interpreted as a compactional feature. The pumice forms are oblate spheres and poorly rounded blocks (Figs 2*b*, 3). Some have a glassy, poorly vesicular outer sheath about 1 cm thick, but this is discontinuous. Many of the blocks have a crude prismatic jointing normal to their margins, most clearly developed in the outer part of the blocks; occasionally this jointing is very well developed (Fig. 2*b*). These features indicate that the giant pumice blocks have been chilled. A small amount of abrasion evidently took place before deposition.

Features such as relative uniformity of pumice size, relative uniformity in thickness (Fig. 1*d*), maximum pumice sizes (Fig. 1*e*) of the bed over practically the entire former lake area, and the lack of bomb sags make the giant pumice block bed sufficiently unique to demand an unusual mechanism of formation. Pyroclastic processes are inadequate to account for these features. The following mechanism (Fig. 1*f*), consistent with all the observed field relationships, envisages porphyritic comendite lava extrusions forming on the floor of a lake. It is common for the carapace of a silicic lava extrusion to be pumiceous, and in this case pumice blocks became detached from the carapace and floated to the lake suface. Eventually, the entire surface was covered with giant blocks of floating pumice, and these blocks became evenly distributed by jostling against one another. Size-sorting occurred as smaller pumice fragments rapidly sank close to the extrusions. Eventually, however, the largest pumice blocks became waterlogged and sank to generate the coarse mode of the bed. In places the bed mantles prophyritic lava, and here rests on a thick apron of less coarse pumice which sank earlier (if, indeed, it floated at all).

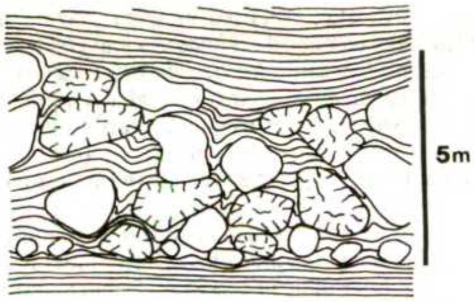


Fig. 3 Field sketch showing disturbed and contorted lake sediments with blocks of giant pumice. Some blocks show cooling joints normal to their margins.

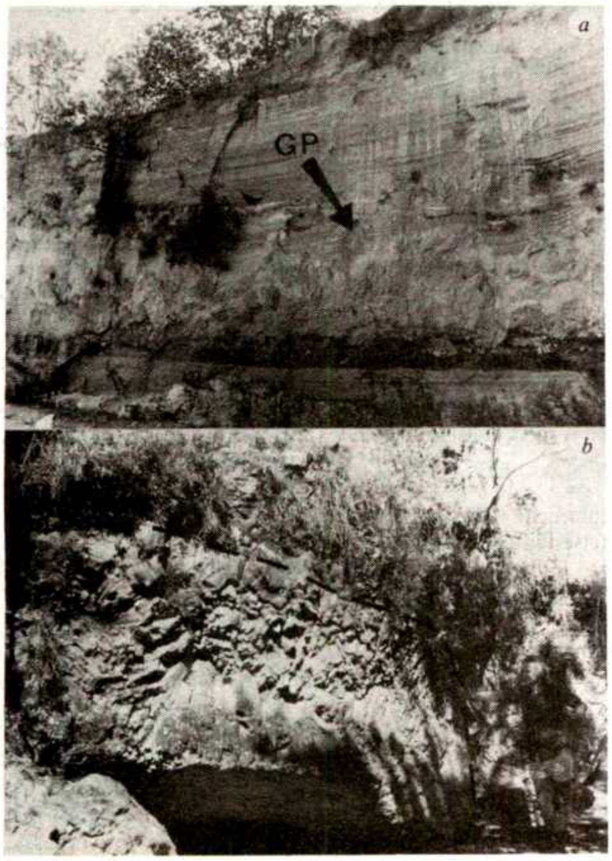


Fig. 2 *a*, Giant pumice block bed (GP), at this location approximately 5 m thick, and laminated lake sediments which are disturbed and contorted. *b*, Giant pumice block; this example has a largest diameter of ~6 m, and also has very well developed prismatic jointing normal to its margins.

Various lines of evidence indicate that the porphyritic lavas were emplaced in water or wet lake sediments. These include, (1) an unusually thick carapace of glassy rocks which is invariably present, (2) particularly well developed jointing (in places the rocks are extensively brecciated), (3) anomalously close spacing of microfractures and a resulting very pervasive perlitic alteration, and (4) an absence of spherulites and lack of post-eruption crystallization in general. In contrast, the aphyric comendite lavas of the volcano (Fig. 1*b*) are thought to have been erupted 'dry'. They have a carapace of fresh or slightly perlitized obsidian interbanded with stony and spherulitic rhyolite, and close-spaced foliation planes which are steeply inclined and intensely contorted.

The origin of the lake depression in which the sediments accumulated is not immediately apparent, because subsequent uplift of the central part of the volcano masks the original relationships (Fig. 1*c*). Geological relationships, however, suggest that the lake occupied a caldera. The sediments extend up to 1 km outside the limit of the rhyolitic ring, which conforms with observations that the topographic rim of a caldera commonly lies outside the ring fault[3] (as marked by lava extrusions or explosive vents). There is, in fact, a low escarpment 1–2 km outside the lava ring on the northern and eastern side (Fig. 1), but this could be a second outer embryo caldera of a nested pair. Resurgent updoming[4] has caused an inversion of topography so that the lake deposits now stand well above any possible shoreline. The updoming took place relatively late in the history of the volcano and did not follow closely on the climatic ignimbrite eruption. At La Primavera this could indicate the arrival of a new pluton beneath the volcano. If so, the apparently youthful age of the updoming indicates that possible future activity of the volcano should be considered seriously[2]. Numerous hot springs within the volcano point to the existence of a hydrothermal system driven by heat from a hot rock or magma body.

We thank NERC for funding field work in 1976 and 1977, and the Director (Ing. Diego A. Cordoba) of the Instituto de Geología, UNAM, for logistic support. B.J.C. and J.V.W. held NERC research studentships during this work.

Received 28 July; accepted 28 October 1980.

1. Wright, J. V. thesis, Univ. London (1979).
2. Walker, G. P. L., Wright, J. V., Clough, B. J. & Booth, B. *Geol. Rundsch.* (in the press).
3. Lipman, R. W. *Geol. Soc. Am. Bull.* **87,** 1397 (1976).
4. Smith, R. L. & Bailey, R. A. *Geol. Soc. Am. Mem.* **116,** 613 (1968).

Microfossil-like objects from the Archaean of Greenland: a cautionary note

D. Bridgwater[*¶], J. H. Allaart[*], J. W. Schopf[†], C. Klein[†#], M. R. Walter[†], E. S. Barghoorn[‡], P. Strother[‡], A. H. Knoll[§] & B. E. Gorman[||]**

[*] Geological Survey of Greenland Øster Voldgade 10, DK-1350 Copenhagen K, Denmark
[†] Precambrian Paleobiology Research Group, Department of Earth and Space Science, University of California, Los Angeles, California 91124
[‡] Department of Biology, Harvard University, Cambridge, Massachusetts 02138
[§] Department of Geology, Oberlin College, Oberlin, Ohio 44074
[||] Department of Geology, The University of Western Ontario, London, Canada N6A 5B7

Recent reports[1–4] have described "yeast-like microfossils" (*Isuasphaera isua* Pflug[2]) in 3,800-Myr old[5] metaquartzites from the Isua supracrustal belt of south-west Greenland. A biogenic interpretation of these objects is inconsistent with the tectonic history of the Isua region[6–8], with the petrology of the metaquartzites, and with the morphology of the microstructures themselves. The putative microfossils are indistinguishable from limonite-stained fluid inclusions: microstructures which are demonstrably inorganic and post-depositional in origin. As such, we contend here that these objects should not be regarded as evidence of early Archaean life forms.

The geological history of the Isua supracrustal belt is summarized in Table 1. Several episodes of deformation, geochemical alteration and amphibolite-facies metamorphism have been recognized, in sharp contrast to slightly younger Archaean strata from which stromatolites[9,10] and apparently *bona fide* microfossils[11–13] have been reported.

Evidence of penetrative deformation is seen throughout the Isua succession where all lithologies record high strains. Few primary sedimentary and igneous structures are preserved apart from the gross lithological layering which allows a rudimentary stratigraphy to be described. Even in fold hinges or inclusions of competent material in a ductile matrix where strain is comparatively low a strong linear fabric is developed and primary features are only recognizable on surfaces normal to the maximum extension. Original discordant structures, such as the contacts between the supracrustal rocks and the intrusive Amîtsoq gneisses have been rotated to near-parallel. The absence of strain markers in the main quartzite horizon (from which the fossil-like microstructures were found), obscures evidence of deformation. Nonetheless, thin layers of magnetite and cummingtonite in the quartzites exhibit extension fabrics similar to those in neighbouring ironstones. There is no evidence that the quartzite horizon escaped the episode(s) of high deformation.

Large-scale geochemical alteration is inferred from rocks interpreted as volcanic in origin. Basic rocks record massive remobilization of alkalis, silica and CO_2 (ref. 14), whereas acid rocks are strongly enriched in K_2O (refs 6, 7) or carbonate minerals. Ultrabasic units (dunites and peridotites) are commonly transformed into schistose assemblages of talc, Mg-amphibole and carbonates. Sedimentary carbonate units display high initial Sr-isotope ratios (see W. Compston in ref. 15) attributed to massive metasomatic changes in constituents.

At least five regional metamorphic events occurred between 3,700 and 1,600 Myr (Table 1). The supracrustal rocks reached amphibolite-facies metamorphism twice: once associated with the emplacement of the Amîtsoq gneisses (~3,700 Myr), and again following the intrusion of the Ameralik dykes (before 3,000 Myr). Efforts to quantify temperatures and pressures attending these metamorphic events have been hampered by the

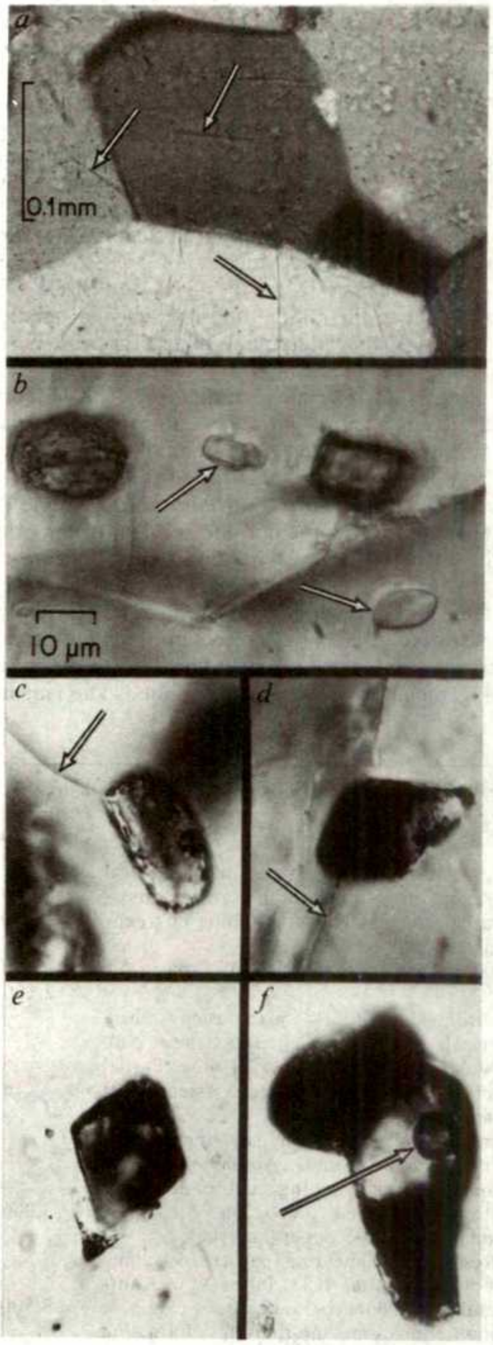


Fig. 1 Optical photomicrographs showing recrystallized, equigranular quartz grains (in plane polarized light; *a*) and non-biogenic, apparently limonite-stained microfossil-like objects (in transmitted white light; *b–f*) in petrographic thin sections of greenish-white quartzite from the 3,800 Myr Isua supracrustal belt of Greenland; *a*, arrows point to fine actinolite needles; *b*, arrows point to clear, ellipsoidal, fluid inclusions; *c*, *d*, arrows point to limonite-stained quartz grain boundaries (also visible in *b*); *f*, arrow points to gas bubble within an elongate, mineral-stained inclusion. *b–f* (scale as in *b*) are from thin section no. 60464, Paleobotanical Collections, Harvard University, Cambridge, Massachusetts.

[¶] Present address: The Geological Museum, Øster Voldgade 5, DK-1350 Copenhagen.
[#] Permanent address: Department of Geology, Indiana University, Bloomington, Indiana 47405.
[**] Present address: Bureau of Mineral Resources, Geology and Geophysics, PO Box 378, Canberra, ACT, Australia 2601.

0028–0836/81/020051—03$01.00

polymetamorphic nature of the rocks and their common retrogression to chlorite-bearing greenschist facies assemblages. High-grade mineral assemblages recognized from different rock types are listed in Table 2. With specific reference to the pelitic units, the assemblage biotite + garnet + quartz + staurolite and the presence of kyanite[16] are consistent with temperatures of 500–600 °C and pressures of 200–500 MPa. Independent evidence of the temperatures of metamorphism is given by the oxygen isotope fractionation in quartz–magnetite pairs of the main quartzite horizon[17]. These ratios indicate a mean (and likely minimum) equilibration temperature of 390 °C, which we interpret as reflecting a late (post-3,000 Myr) event following regional amphibolite-facies metamorphism.

The 'fossiliferous' rock, specimen no. 2377, is a relatively coarse-grained, sugary, greenish-white quartzite. It is composed of quartz (95%), actinolite (4%) and talc (1%). The quartz is equigranular and totally recrystallized, with most grains clustering in the 0.2–0.3 mm size class. Neither the presence of "chalcedony" in the rock[1,3,4], nor its supposed "cherty" texture[1,2,4]. could be confirmed. Much of the actinolite $(Ca_{1.44}Na_{0.04}Mn_{0.07}Fe_{0.45})$ $(Mg_{3.79}Fe_{1.13}Al_{0.07})$ Si_8O_{23} is of a grain size similar to the quartz, and fine acicular needles can be found within many individual quartz grains (Fig. 1*a*). The talc $(Mg_{5.19}Fe_{0.87})$ $(Si_{7.90}Al_{0.07})O_{22}$ is seen as fine "beards" along quartz grain boundaries. The mineralogy of the quartzite does not contradict our conclusion that the quartzite unit was subjected to amphibolite-facies regional metamorphism.

Evidence of high grade metamorphism and extreme deformation does not entirely preclude the survival of microfossils, although the chances of preserving intact such delicate carbonaceous bodies as those described following complete recrystallization of their host rock seem to be remote. The morphology of the microfossil-like objects in thin section (including specimen no. 2377 from which *Isuasphaera* was first described[2]) convinces us that a biogenic interpretation is unfounded. The Isua microfossil-like objects are reported to "consist predominantly of carbon . . . characterized by a high degree of bitumunization"[1] preserved "partly in a charred condition, partly close to a final stage of graphitization"[4]; they have shapes ranging from ellipsoidal to elongate with "individuals . . . found in the assemblages which resemble a biscuit, a dumb-bell, or a pair of tears"[1]. The objects range in size from 10 to 40 μm (with "mature cells" said to occur towards the upper end of the size range[3]), and they are described as being "hollow" with their interiors "partly filled with organic matter or fine-grained pyrite" which is "dark brownish" in colour[1,3]. They are also reported to exhibit structures interpreted as "cell walls", "cell vacuoles", "gas vacuoles", "multilaminate sheaths", "budding structures", "bud scars", "cross walls", "remnants of protoplasm" and evidence of "filamental growth"[1,3]. By comparison with illustrations of dried, carbol fuchsin-stained yeast cells, and on the basis of the features noted above, these objects have been interpreted as representing carbonaceous cellularly intact remnants of "yeast-like organisms", possibly of "an evolutionary level which is far below that of yeasts of today" but which "may represent a half-way line between a microsphere-like protobiont and subsequent evolution"[1,3].

Table 1 Simplified table of events in the Isua area

Events	Time (Myr BP)
10. Resetting of K–Ar and Rb–Sr mineral ages[25] local granite intrusion[26] and static thermal metamorphism to approximately 300°, faulting in greenschist facies conditions.	1,600
9. Intrusion of E–W, N–S and NE–SW Proterozoic dykes.	2,100–
8. Transcurrent faulting juxtaposing Isua supracrustals and Amîsoq gneisses against Nûk (~3,000 Myr) gneisses in NW of area. Partial retrogression of amphibolite-facies assemblages in supracrustals.	
7. Deformation with local strong shearing concentrated in supracrustal units. Amphibolite-facies metamorphism and strong deformation of Ameralik dykes within supracrustals.	3,000–
6. Intrusion of E–W, N–S Ameralik dykes.	
5. Retrogression and shearing particularly along contact of supracrustal rocks. Intrusion of Amîtsoq pegmatites and leucogranites.	3,500–
4. Amphibolite-facies metamorphism, deformation to produce major stretching fabric in supracrustals.	
3. Intrusion of Amîtsoq tonalitic and granodioritic gneisses.	3,700–
2. Deformation of supracrustals. Fabric locally preserved, mineral assemblages commonly retrogressed during later events.	
1. Deposition of Isua sediments, volcanism, post-depositional hydrothermal activity.	3,800–

Ages quoted are rounded to the nearest 100 Myr. Based on published and unpublished Rb–Sr, Sm–Nd, Pb–Pb whole-rock, Pb–Pb, U/Pb, Rb–Sr and K–Ar mineral determinations. See refs 5–8 for original refs.

Table 2 Amphibolite-facies mineral assemblages* in the Isua supracrustal rocks

Rock type	Assemblage
Ironstones (and Mg-rich analogues)	Quartz, magnetite, grunerite/cummingtonite
Basic volcanic rocks	Hornblende, plagioclase (andesine–labradorite), quartz, magnetite, sphene (diopside)
Mg-rich and ultrabasic volcanic rocks	Talc, anthophyllite, carbonate minerals (quartz, plagioclase)
Carbonate rocks and calc-silicates	Diopside, anthophyllite, Fe-carbonate, calcite
Pelites	Biotite, quartz, garnet, magnetite (staurolite, kyanite)

* Minerals in parentheses are only sporadically preserved.

Our thin sections contain microstructures which conform in size, shape and all other optically detectable properties to the putative microfossils previously described, leaving no doubt that our material is identical to that previously interpreted as biogenic. We find, however, that the populations of these objects span a decidedly broader size-range than that previously reported: measurements of 250 such objects (in three thin sections) show that they range in length between 2 and 115 μm, with a mean length of ~23 μm (and a variance, s^2, of 38.2).

Two principal categories of fossil-like objects occur in the thin sections: (1) clear, spherical to ellipsoidal structures (Fig. 1*b*) occur only within quartz grains (rather than at grain boundaries); these structures are comparable to fluid inclusions: inorganic, metamorphically produced micro-structures that are of well known and relatively common occurrence in other, similarly altered metasediments[18]; (2) yellowish to reddish-brown (apparently limonite-stained) structures that exhibit a continuum of shape from spheroidal to ellipsoidal to subangular to completely euhedral (Fig. 1*b–f*).

With the apparent exception of the occurrence of "multilaminate sheaths" (of which we find no evidence) members of this second category are identical to forms described as *Isuasphaera*[1–4]. However, the population includes rhombohedral (Fig. 1*e*) and obviously crystalline individuals (see also ref. 1 Fig. 3, parts 2, 4, 11, 12) which do not exhibit the pattern of morphological consistency characteristic of fossilized microorganisms[19]. Furthermore the objects show a marked concentration at quartz grain boundaries (Fig. 1*c*, *d*), a feature commonly observed in fluid inclusions in recrystallized granites[20] and not likely to be the result of a primary sedimentological distribution. The reported "budding", which has previously been attributed to a biological origin, we interpret as a 'necking-off' process commonly occurring during recrystallization of multiphase inclusions[21], while the reported "gas vacuoles" are bubbles trapped within the objects during their formation and recrystallization[18]. Other suggested biological features (bud scars, cross walls, and so on) are more consistent with an inorganic origin, for example, the staining of the surfaces of fluid inclusions and negative crystals by secondarily mobilized limonitic iron oxides which were also deposited along grain boundaries. Many of the quartz grain boundaries are, in fact, stained with limonite and are indistinguishable in colour and texture from the coating that outlines the fossil-like objects (see ref. 4, plate 2, Figs 2, 5, 6).

Thus consideration of the geological, petrological and morphological evidence leads us to conclude that the fossil-like microstructures (*Isuasphaera isua*) described from the 3,800-Myr old Isua supracrustal rocks are in fact non-biogenic artefacts of inorganic, post-depositional origin. Their reported "yeast-like" morphology and fine structure are inconsistent with the tectonic history of the rocks in which they occur and are at variance with the remainder of the Precambrain fossil record as now known[22]. The occurrence of hydrocarbons within these microstructures[3], if confirmed, would be equally consistent with both a biological and non-biological origin of these compounds. We believe that the microstructures are non-biogenic, and while graphitic carbonaceous matter occurs in the Isua metasediments[4,23,24] it remains to be established whether this is biogenic or abiogenic in nature and to what extent it may be of postdepositional rather than solely syngenetic origin.

We thank M. Schidlowski for providing sample no. 2377 and P. Appel and C. Walters for additional material, and W. S. Fyfe, G. Perry, R. Dymek, R. Kerrich and J. D. Meloche for information and discussion. Mapping at Isua (D.B. and J.H.A.) is part of the programme of the Geological Survey of Greenland. Field work was supported by NATO research grant no. 949 and by a NSERC (Canada) grant to W. S. Fyfe (B.E.G.). We thank other members of the NATO supported research group[8] for discussion. Laboratory work was supported by NSF grants 77-22518 and 79-21777 and by NASA grants NGR 05-007-407 and NSG 7489 (to J.W.S.); by NASA grant NGL 22-007-067 and by NSF grant EAR 78-24237 (to E.S.B.) by NSF grant EAR 76-11740 (to C.K.).

Received 2 May, accepted 10 November 1980.

1. Pflug, H. D. *Naturwissenschaften* **65,** 611–615 (1978).
2. Pflug, H. D. *Oberhess. Naturw. Z.* **44,** 131–145 (1978).
3. Pflug, H. D. & Jaeschke-Boyer, H. *Nature* **280,** 483–486 (1979).
4. Pflug, H. D., Jaeschke-Boyer, H. & Sattler, E. L. *Microsc. Acta* **82,** 255–266 (1979).
5. Moorbath, S., O'Nions, R. K. & Pankhurst, R. J. *Nature* **245,** 138–139 (1973).
6. Bridgwater, D., Keto, L., McGregor, V. R. & Myers, J. S. in *Geology of Greenland* (eds Escher, A. & Watt, W. S.) 18–75 (Geological Survey of Greenland, 1976).
7. Allaart, J. H. in *The Early History of the Earth* (ed. Windley, B. F.) 177–189 (Wiley, New York, 1976).
8. Bridgwater, D. *et al. Rapp. Grønland geol. Unders.* **95,** 66–71 (1979).
9. Walter, M. R., Buick, R. & Dunlop, J. S. R. *Nature* **284,** 443–445 (1980).
10. Lowe, D. *Nature* **284,** 441–443 (1980).
11. Schopf, J. W. & Barghoorn, E. S. *Science* **156,** 508–512 (1967).
12. Muir, M. D. & Grant, P. R. in *The Early History of the Earth* (ed. Windley, B. F.) 595–604 (Wiley, New York, 1976).
13. Knoll, A. H. & Barghoorn, E. S. *Science* **198,** 396–398 (1977).
14. Gill, R. C. O., Bridgwater, D. & Allaart, J. H. *Spec. Publ. geol. Soc. Aust.* (in the press).
15. Apple, P. W. U. *Precamb. Res.* **11,** 83 (1980).
16. Boak, J. L. & Dymek, R. F. *Geol. Soc. Am. Abstr.* (in the press).
17. Perry, E. C., Ahmad, S. N. & Swulius, T. M. *J. Geol.* **86,** 223–239 (1978).
18. Deicha, G. *Les Lacunes des Cristaux et leurs Inclusions Fluides. Signification dans la Genèse des Gîtes Minéraux et des Roches* (Masson et Cie, Paris, 1955).
19. Schopf, J. W. *Origins of Life* **7,** 19–36 (1976).
20. Wilkins, R. W. T. & Barkas, J. P. *Contr. Miner. Petrol.* **65,** 293–299 (1978).
21. Roedder, E. *Fluid Inclusion Res. Proc. Coffi*, 4 (1968).
22. Schopf, J. W. *A. Rev. Earth planet. Sci.* **3,** 213–249 (1975).
23. Nagy, B., Zumberge, J. E. & Nagy, L. A. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1206–1209 (1975).
24. Schidlowski, M., Appel. P. W. U., Eichmann, R. & Junge, C. E. *Geochim. cosmochim. Acta* **43,** 189–199 (1979).
25. Pankhurst, R. J., Moorbath, S., Rex, D. C. & Turner, G. *Earth planet. Sci. Lett.* **20,** 157–170 (1973).
26. Kalsbeek, F., Bridgwater, D. & Boak, J. *Rapp. Grønlands geol. Unders.* **100,** 73–75 (1980).

Amino acids and hydrocarbons $\sim$3,800-Myr old in the Isua Rocks, southwestern Greenland

Bartholomew Nagy*, Michael H. Engel*‡, John E. Zumberge*§, Hiroshi Ogino* & Sai Y. Chang†||

* Laboratory of Organic Geochemistry, Department of Geosciences, The University of Arizona, Tucson, Arizona 85721
† Department of Pharmacology, The University of Arizona, Tucson, Arizona 85721

The Isua supracrustal belt consists of metamorphic rocks, including metamorphosed sediments, and are the oldest known rocks on Earth. It has been postulated[1,2] that hydrocarbons recently detected in these rocks are remnants of organisms that lived $\sim$3,800 Myr ago. This popular interpretation is inconsistent with the high-temperature history of the Isua rocks. Various types of hydrocarbons were liberated from these rocks after repeated solvent extraction by matrix dissolution or pyrolysis. However, the Isua rocks have been metamorphosed to the upper greenschist or amphibolite facies (400–600 °C); the metamorphic episodes lasted $>10^6$ yr. Consequently, it is very important that the origins of the key organic constituents be established. In the study reported here both hydrocarbons and amino acids were detected in Isua rock samples. In addition to the common biological amino acids, geologically unstable amino acids and biologically uncommon ones were detected (such as sarcosine). The extent of amino acid racemization suggests the apparent continuous diffusion of biochemicals into the Isua rocks from encrusting lichens since the end of the last ice age. Also, the cold-temperature history (which retards racemization) suggests that the amino acids may be modern to a few tens of thousands of years old. *n*-Alkanes lacked odd/even carbon chain preference which may indicate some antiquity. However, the hydrocarbons also resemble a recent petroleum distillate fraction. Laboratory experiments, supported by kinetic and thermodynamic calculations, show that amino acids and aromatic and saturated aliphatic hydrocarbons, including pristane, could not have survived the known metamorphic history of the Isua rocks.

Samples from the Isua banded iron formation ($\sim$150 km north-east of Godthåb) were analysed. Sample 155782 was an interior portion of the banded ironstone ($\sim$50 g) that contained a limonite-filled microcrack. Samples 3015, 3016, 3022 and 3028 ($\sim$50 g each) had weathered surfaces. Porosity, air and water permeabilities of samples 3015 and 3016 were 2.3%, 1.0×10^{-2} mdarcies and 1.2×10^{-6} mdarcies; and 2.2%, 4.6×10^{-1} mdarcies and 1.3×10^{-1} mdarcies, respectively[3]. Isua sample 3015 was less extensively weathered than Isua 3016 and thus had a lower water permeability.

To minimize laboratory contamination, all glassware was acid cleaned, solvents and acids were redistilled from reagent or spectral grade, water was triple distilled in glass and so on[4,5]. The rock surfaces were drilled off and the samples pulverized. Excessive ball milling (>20 h) was avoided because it caused some racemization of amino acids. The pulverized samples and corresponding blanks were refluxed with water for 8 h. The filtrate residues were hydrolysed for 24 h at 100 °C in 6 M HCl. The samples were then desalted by ion exchange chromatography[4,5] and the *N*-PFP-(+)-2-butyl amino acid diastereomers were prepared[4,5] and analysed for their D/L amino acid ratios by gas chromatography. Amino acids in the water extract of sample 3028 were analysed by combined gas chromatography–chemical ionization mass spectrometry (GC–CIMS). Part of the water extract of Isua sample 3016 was analysed by an amino acid analyser.

Present addresses: ‡Carnegie Institution of Washington, Geophysical Laboratory, 2801 Upton St. NW, Washington, DC 20008; §Cities Service Co., Box 50408, Tulsa, Oklahoma 74150; ||McNeil Laboratories, 500 Office Center Drive, Fort Washington, Pennsylvania 19034.

0028-0836/81/020053—04$01.00

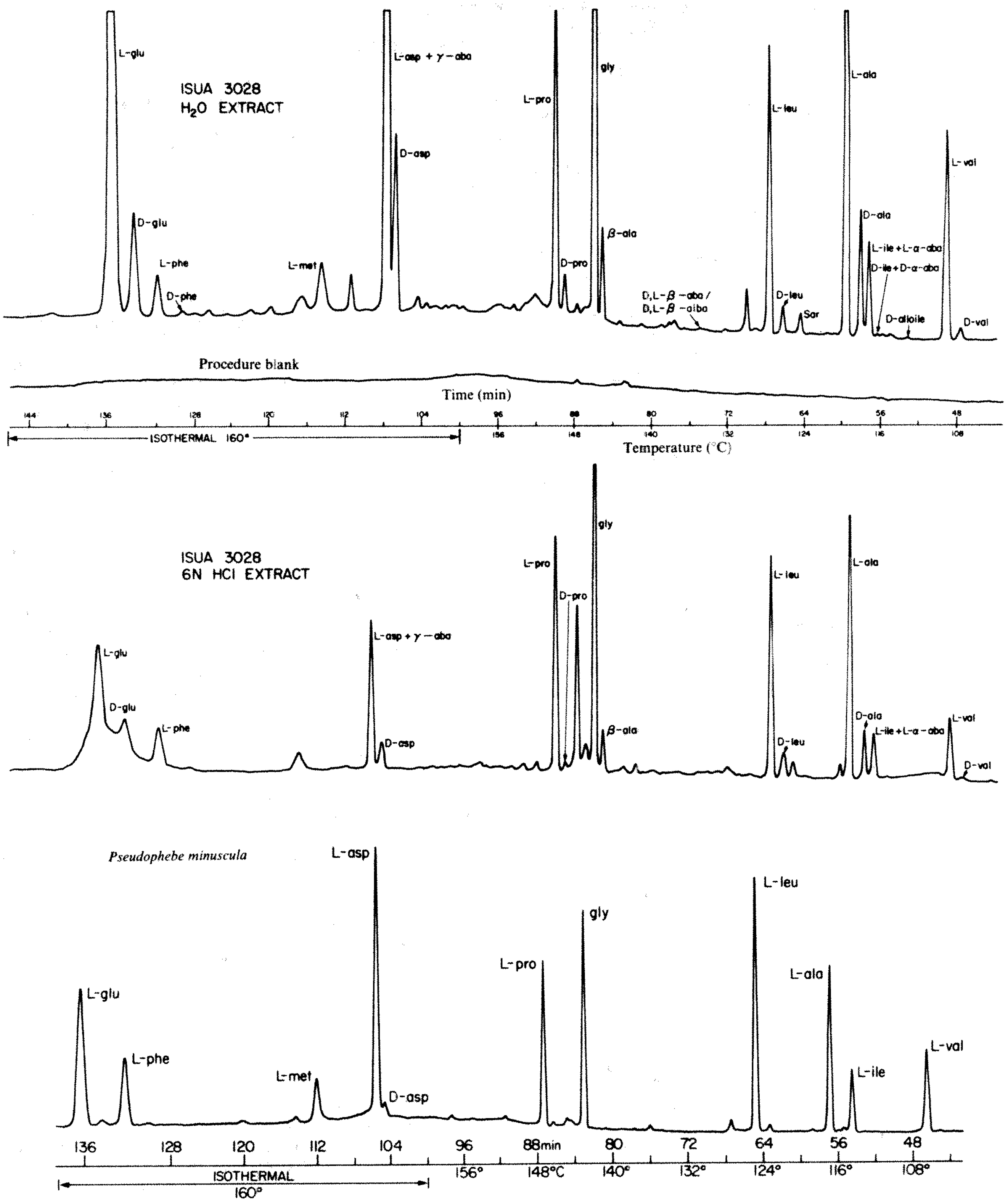


Fig. 1 Gas chromatograms of the water extract, procedure blank and HCl acid extract of Isua sample 3028 and of the 6M HCl extract of the lichen *Pseudophebe minuscula*. A Perkin-Elmer model F-11 gas chromatograph coupled to an Infotronics CRS-208 integrator was used with a Carbowax-20M 61m × 0.5 mm i.d. WCOT nickel capillary column. The He flow rate was 11.5 ml min^{-1}. The inlet temperature was 150 °C. The oven programme was 80 °C isothermal for 20 min; then 80 °C to 160 °C at 1 °C min^{-1}. The programmes were finished isothermal at 160 °C. In addition to co-injection of amino acid standards and retention times, all amino acid identifications in the water extract of sample 3028 were confirmed by GC–CIMS. A Finnigan 9500 gas chromatograph coupled to a Finnigan 3300 mass spectrometer (quadrupole) equipped with a 6110 series data system was used. CH_4 served as both carrier gas and reactant for chemical ionization. A Carbowax-20M 46 m × 0.5 mm i.d. SCOT stainless-steel capillary column was used for GC–CIMS. The carrier gas flow rate was ~8.5 ml min^{-1}. The oven programme was 90 °C to 150 °C at 2 °C min^{-1}. The programme was finished isothermal at 150 °C. The ionization chamber temperature and pressure were 150 °C and 1.0 Torr, respectively. The emission current was 1.0 mA. 150 eV was used to generate the methane ions. Valine, val; alloisoleucine, alloile; α-amino-*n*-butyric acid, α-aba; isoleucine, ile; alanine, ala; sarcosine, sar; leucine, leu; β-amino-*n*-butyric acid, β-aba; β-aminoisobutyric acid, β-aiba; β-alanine, β-ala; glycine, gly; proline, pro; γ-amino-*n*-butyric acid, γ-aba; aspartic acid, asp; methionine, met; phenylalanine, phe; glutamic acid, glu.

Table 1 Isua rocks and lichen species amino acid D/L ratios

Sample	Extraction	Concentration†	Glu	Phe	Asp	Pro	Leu	Ala	Ile	Val
3016	H_2O	26.6	0.269	‡	0.177	0.192	0.115	0.220	0.053	0.073
3016	HCl	49.2	0.158	‡	0.127	0.066	0.023	0.088	‡	§
3022	H_2O	8.0	0.141	‡	0.162	0.181	0.061	0.108§	‡	‡
3022	HCl	13.4	0.076‡	‡	0.124	0.054	§	§	‡	‡
3028	H_2O	62.2	0.186	0.064	0.150	0.127	0.094	0.206	0.028	0.068
3028	HCl	7.9‖	§	§	0.194§	0.045	0.140§	0.174	‡	§
155782*	H_2O	1.3	0.087‡	‡	0.178	0.229	‡	0.340	‡	‡
155782*	HCl	3.6	‡	‡	‡	‡	‡	‡	‡	‡
Rhizocarpon sp.	HCl	4.5	0.034	0.023	0.059	0.052	0.032	0.022	‡	0.027
Dermatocarpon sp.	HCl	104.0	0.033	0.024	0.050	0.035	0.027	0.029	‡	0.027
Pseudophebe minuscula										
Free fraction	HCl	0.1	0.025	‡	‡	0.035	‡	‡	‡	‡
Bound fraction	HCl	1.6	0.028	‡	0.048	0.023	0.029	0.026	‡	0.017

Approximately 2.0% of the D-amino acids used to obtain the D/L ratios may be attributed to (−)-2-butanol impurities and racemization during 6M HCl hydrolysis. Average instrumental error of three gas chromatographic determinations for each D/L ratio = ±0.011. Asp D/L ratios could not be corrected for minor γ-aba interference.

* This rock specimen is another portion of the sample on which the original Pb/Pb date of the Isua rocks was determined[30]. This age has been confirmed by other methods, see ref. 31.

† These values are total concentrations of the amino acids detected in the samples. The Isua extracts = nmol g^{-1}; the lichen extracts = nmol $mg^{-1} \times 10^{-2}$. The experimental error for the amino acid concentration is ±20%.

‡ D peak and/or L peak too small to determine D/L ratios. The two values reported are based on only single gas chromatographic determinations.

§ Interference of D peak and/or L peak.

‖ Glu and Phe were not included in this value because of gas-chromatographic peak interference.

The water-extracted rock residues were digested for 24 h at 100 °C in 6M HCl. The HCl solutions were analysed by the same procedure as above. During the HCl digestion the possibility arose that release of ferric iron or other substances may have oxidized the amino acids. To test this, an 8.1-g portion of sample 155782 was mixed with L-norvaline (1.7 mg) and L-norleucine (1.9 mg), two amino acids not originally detected in 155782. The mixture was digested and analysed as before. Approximately 75% of these amino acids were recovered after HCl digestion. No racemization was noted. 6M HCl extracts of two lichen species from the surface of sample 3016 (*Rhizocarpon* sp., *Dermatocarpon* sp.) and one lichen species from 3028 (*Pseudophebe minuscula*) were analysed for amino acids. Before acid hydrolysis *P. minuscula* was sonicated for 30 min with 1.5 M HCl to liberate non-bound amino acids and/or peptides. This free fraction was hydrolysed separately.

Figure 1 shows the results of the analysis of sample 3028 and of the 6M HCl extract of *P. minuscula*. Notable similarities exist between the relative abundances of amino acids in the lichens and in the Isua rocks. Glutamic acid, aspartic acid and glycine, the most abundant amino acids in the three lichen species analysed, were also the most abundant amino acids in the Isua rocks. Co-injection and GC–CIMS confirmed the presence of small quantities of four uncommon non-protein amino acids (α-amino-*n*-butyric acid, β-amino-*n*-butyric acid and/or β-aminoisobutyric acid, γ-amino-*n*-butyric acid and sarcosine) in the 3028 water extract. Three of these amino acids have been detected in biological systems[6]; β-amino-*n*-butyric acid and β-aminoisobutyric acid can also be derived biologically. α-Amino-*n*-butyric acid and γ-amino-*n*-butyric acid can evolve by the decomposition of threonine and glutamic acid, respectively[7–9].

D/L ratios and the abundances of the amino acids from the Isua rocks and lichens were determined by gas chromatography and are listed in Table 1. Part of the Isua 3016 water extract that was analysed on an amino acid analyser revealed the presence of threonine, serine, tyrosine, histidine, arginine and lysine (P. E. Hare, personal communication), in addition to the amino acids listed for sample 3016 in Table 1. In temperate environments containing moisture, syngenetic amino acids should be substantially racemized (or epimerized) after ~a few hundred thousand to a few million years[10–13]. The amino acids found in the ~3,800-Myr-old Isua metamorphosed sedimentary rocks were not racemic. The high L-amino acid abundances introduced by modern lichens may have lowered the D/L ratios of older amino acids in the Isua rocks. Amino acids extracted from the Isua rocks with 6M HCl seem to be more tightly bound to the organic and/or inorganic rock matrices than the water-extractable amino acids; they yielded lower D/L ratios. This is not surprising considering recent reports on the adsorbtion of amino acids by organic and inorganic rock constituents and their concomitant effects on amino acid racemization rates[14–17].

After deposition, the Isua rocks experienced episodes of metamorphism. During the metamorphic events temperatures were 400–600 °C (refs 18, 19). Elevated temperature experiments have demonstrated that amino acids and peptides rapidly decompose when exposed to temperatures within the range of the Isua metamorphism[20,21]. Serine, threonine and tyrosine are particularly unstable at elevated temperatures and in geological environments[22,23]. Their presence further precludes an elevated temperature history for the amino acids in the Isua rocks.

To test the possibility of diffusion of amino acids from the lichens into the rocks, fragments of samples 3016 and 3015 were

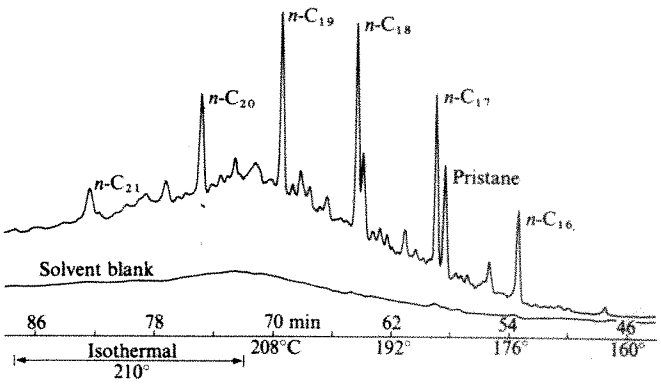


Fig. 2 Gas chromatrograms of the procedure blank and the saturated hydrocarbons in sample 3016. The alkane distribution pattern may suggest some antiquity and/or contribution from a petroleum distillate fraction. For gas chromatography, an Apiezon L 15 m × 0.5 mm i.d., WCOT stainless-steel column was used. The GC programme was 80 °C isothermal for 6 min; then 80–210 °C at 2 °C min^{-1}. The He carrier gas flow rate was 3.5 ml min^{-1}. The inlet temperature was 195 °C. The chromatograms shown above were obtained using a Perkin Elmer F-11 gas chromatograph. In addition to co-injection of standards, the major hydrocarbon components in samples 3016 and 3028 were identified by GC–MS.

placed in aqueous solutions of D-amino acids for 1,321 h at 36 °C. After soaking, the concentrations of the D-amino acids in the rock interiors were 1 or 2 orders of magnitude higher than before exposure. Some amino acids present in the Isua rocks were probably introduced by encrusting lichens. However, the lichens chosen for analysis in this study were apparently not the direct source of the less common amino acids such as α-amino-n-butyric acid, β-amino-n-butyric acid/β-aminoisobutyric acid and sarcosine which were present in low concentrations in the Isua rocks.

Comparisons of the Isua rock amino acid D/L ratios with amino acid D/L ratios reported for samples from other regions that have experienced comparable cold-temperature histories, such as the late Pleistocene fossil shells from the eastern Canadian Arctic[24] and shells from a Carribean core (bottom water = 4 °C)[25], indicate that some of the Isua amino acids may be several tens of thousands of years in age; although comparisons of amino acid D/L ratios in rocks with fossil shells are tenuous. Comparisons of the amino acid D/L ratios of the Isua rocks with those reported from clay-rich (low-carbonate) sediments from Leg 15, Site 148, of the Deep Sea Drilling Project[26] and with amino acid D/L ratios reported for siliceous oozes from a core taken from the southern Atlantic Ocean (lat 56°34.7′ S, long 20°17.2′ W; bottom water, −0.2 °C)[27] further suggest that none of the amino acids in the Isua rocks are older than a few tens of thousands to possibly a few hundreds of thousands of years.

Interior portions of Isua samples 3016 and 3028 were also analysed for saturated hydrocarbons. The rock powders were first extracted with H_2O at 100 °C followed by $CHCl_3/CH_3OH$ (87/13, v/v) at room temperature. Two alumina columns (Woelm Neutral, Waters Associates) were prewashed with distilled n-hexane. Procedure blanks and sample extracts were eluted from the alumina columns with n-hexane. The samples and blanks were analysed by gas chromatography and GC–MS.

Gas chromatograms of the procedure blank and hydrocarbon extract of sample 3016 are shown in Fig. 2. A similar hydrocarbon envelope was detected from sample 3028. The aliphatic hydrocarbons in samples 3016 and 3028 ranged from n-C_{16} to n-C_{21}. Pristane was present in both samples. No odd carbon number preference was found for these Isua hydrocarbon extracts, indicating that these hydrocarbons were old or were introduced more recently from some organisms lacking odd carbon preference. Alternatively, they may have been derived from older materials brought into contact with the rocks by glacial action or by contamination with petroleum products.

It may clarify the origin of the hydrocarbons found in the Isua rocks[1] if pertinent laboratory experiments and related kinetic and thermodynamic calculations are considered. Elevated temperature experiments have shown that, at 375, 400 and 600 °C, the temperature range of the Isua metamorphic episodes, straight chain aliphatic hydrocarbons decompose in ~$6\frac{1}{4}$ days, ~5 days and in ~10 s, respectively[28,29]. In experiments conducted in this laboratory, pristane (10 μl); pristane (7 μl) and graphite (10.2 mg); and pristane (6 μl) and a Santa Barbara Basin Recent marine sediment (3.7 mg) were heated in glass tubes under He at 520 °C for 2,924, 1,891 and 2,107 h, respectively. Next, the samples were extracted with n-hexane by ultrasonication and analysed by GC–MS. Pristane was destroyed in all three samples. ΔG for the decomposition of nonadecane ($C_{19}H_{40} \rightarrow 19C + 20H_2$) at 427 °C is −212.68 kcal mol^{-1}, (S. L. Miller, personal communication). ΔG for the decomposition of naphthalene ($C_{10}H_8 \rightarrow 10C + 4H_2$) at 427 °C is −79.45 kcal mol^{-1}, and for biphenyl ($C_{12}H_{10} \rightarrow 12C + 5H_2$) at 427 °C is −101.63 kcal mol^{-1}. This indicates that even aromatic hydrocarbons, if initially present, may not have survived prolonged Isua metamorphic episodes. Thus the limited thermal stabilities of alkanes, these aromatic hydrocarbons, amino acids, and the low amino acid D/L ratios, probably preclude a very great age.

We thank S. Moorbath for a rock sample and for reviewing the manuscript, and M. Schidlowski for rock samples and helpful comments, T. Nash for identifying the lichen species. P. E. Hare for help with the amino acid analyser analyses, E. S. Barghoorn for helpful comments and J. L. Bada and G. Eglinton for reviewing the manuscript. This research was supported by NASA grant NGR 03-002-171.

Received 21 July; accepted 8 October 1980.

1. Walters, C. *et al. Abstr. 178th Am. chem. Soc. A. Meet.*, Part 1, Washington, DC Geoc. 5 (1979).
2. *Chem. Engng News* **75**, No. 38, 22–23 (1979).
3. Core Laboratories, *Rep. 79072-1* (1979).
4. Engel, M. H., Zumberge, J. E. & Nagy, B. *Analyt. Biochem.* **82**, 415–422 (1977).
5. Zumberge, J. E., Engel, M. H. & Nagy, B. in *Biogeochemistry of Amino Acids* (eds, Hare, P. E., Hoering, T. C., King, K. Jr) 503–525 (Wiley, New York, 1980).
6. Meister, A. *Biochemistry of the Amino Acids* Vol. 1 (Academic, New York, 1965).
7. Schroeder, R. A. *Geol. Soc. Am Abstr. Prog.* **9**, 1163 (1977).
8. Bada, J. L. & Shou, M.-Y., *Geol. Soc. Am. Abstr. Prog.* **9**, 885–886 (1977).
9. Povoledo, D. & Vallentyne, J. R. *Geochim. cosmochim. Acta* **28**, 731–734 (1964).
10. Abelson, P. H. & Hare, P. E. *Yb. Carnegie Instn Wash.* **67**, 208–210 (1969).
11. Bada, J. L. *Earth planet. Sci. Lett.* **15**, 223–231 (1972).
12. Kvenvolden, K. A. *Space Life Sci.* **4**, 60–68 (1973).
13. Dungworth, G. *Chem. Geol.* **17**, 135–153 (1976).
14. Carter, P. W. *Geochim. cosmochim. Acta* **42**, 1239–1242 (1978).
15. Carter, P. W. & Mitterer, R. M. *Geochim. cosmochim. Acta* **42**, 1231–1238 (1978).
16. Bondy, S. C. & Harrington, M. E. *Science* **203**, 1243–1244 (1979).
17. Hare, P. E. & Hoering, T. C. *Yb. Carnegie Instn Wash.* **76**, 625–631 (1977).
18. Nagy, B., Nagy, L. A., Zumberge, J. E., Sklarew, D. S. & Anderson, P. *Precambr. Res.* **5**, 109–120 (1977).
19. Appel, P. W. U. *Precambr. Res.* **11**, 73–87 (1980).
20. Vallentyne, J. R. *Geochim. cosmochim. Acta* **28**, 157–188 (1964).
21. Simmonds, P. G., Shulman, G. P. & Stembridge, C. H., *J. chromatogr. Sci.* **7**, 36–41 (1969).
22. Hare, P. E. *Yb. Carnegie Instn Wash.* **75**, 801–806 (1976).
23. Williams, K. M. & Smith, G. G. *Origins of Life* **8**, 91–144 (1978).
24. Miller, G. H. & Hare, P. E. *Yb. Carnegie Instn Wash.* **74**, 612–617 (1975).
25. Kvenvolden, K. A., Peterson, E., Wehmiller, J. & Hare, P. E. *Geochim. cosmochim. Acta* **37**, 2215–2225 (1973).
26. Bada, J. L. & Man, E. H. *Earth Sci. Rev.* (in the Press).
27. Warnke, D. A., Blunt, D. J. & Pollock, G. E. in *Biogeochemistry of Amino Acids* (eds Hare, P. E., Hoering, T. C. & King, K., Jr) 183–189, (Wiley, New York, 1980).
28. Henderson, W., Eglinton, G., Simmonds, P. & Lovelock, J. E. *Nature* **219**, 1012–1016 (1968).
29. Abelson, P. H. in *Researches in Geochemistry* (ed. Abelson, P. H.) 79–103 (Wiley, New York, 1959).
30. Moorbath, S., O'Nions, R. K. & Pankhurst, R. J. *Nature* **245**, 138–139 (1973).
31. Hamilton, P. J., O'Nions, R. K., Evensen, N. M., Bridgwater, D. & Allaart, J. H. *Nature* **272**, 41–43 (1978).

Preclassic lowland maize from Cuello, Belize

Charles H. Miksicek*, Robert McK. Bird†,
Barbara Pickersgill‡, Sara Donaghey§,
Juliette Cartwright¶ & Norman Hammond§

* Arizona State Museum and Department of Arid Lands Studies,
University of Arizona, Tucson, Arizona 85721
† Institute for the Study of Plants, Food and Man, PO Box 9983,
Kirkwood, Missouri 63122
‡ Department of Agricultural Botany, University of Reading,
Reading RG6 2AS, UK
§ Archeological Research Program, Douglass College, Rutgers
University, New Brunswick, New Jersey 08903
¶ Department of Anthropology, University of Texas, Austin,
Texas 78712

In spite of recent progress towards an understanding of prehispanic Maya agriculture, relatively little is known about plants grown prehistorically in the Petén, Belize and the Yucatan Peninsula. Little plant material has been reported from Maya sites, and almost no detailed archaeobotanical analyses have been done. The Cuello site (18°05′N, 88°35′W) in northern Belize is unique in having numerous intact floors of very early ceremonial structures which separate plant remains into clearly defined units. Flotation of soil samples has yielded a wide range of carbonized plant material including many maize cupules (cob fragments) and kernels. As we report here, these can be classified into several types which seem to relate to a range of archaeological types of maize found elsewhere.

0028-0836/81/020056—04$01.00

Table 1 Preclassic Cuello maize data

No.	Row no.	Cupule width	Cupule length	Cupule wing width	Rachis diameter†	Cupule thickness	Cupule depth	% Of all maize from each phase Swasey	Mamom	Chicanel
Swasey type 1 (2000 b.c.–a.d. 250)										
Mean 12	12.8	2.4	1.6	0.8	5.2	1.7	1.2	31	6	2
σ	3.6	0.24	0.33	0.20		0.68	0.29			
Swasey type 2 (2000 b.c.–a.d. 250)										
Mean 49	12.9	2.9	1.7	0.9	6.2	2.0	1.1	42	35	28
σ	2.4	0.23	0.24	0.23		0.36	0.48			
Swasey type 3 (2000 b.c.–a.d. 250)										
Mean 44	11.5	3.6	2.1	1.2	7.2	2.3	1.2	27	41	27
σ	2.1	0.27	0.32	0.37		0.55	0.44			
Mamom type 1 (1000 b.c.–a.d. 250)										
Mean 13	11.4	4.6	2.2	1.3	8.6	2.9	1.3		14	10
σ	2.6	0.15	0.27	0.21		0.21	0.21			
Mamom type 2 (1000 b.c.–a.d. 250)										
Mean 7	10.0	5.9	2.5	1.9	10.1	3.0	1.4		4	8
σ	1.3	0.49	0.50	0.29		0.57	0.40			
Chicanel type 1 (400 b.c.–a.d. 250)										
Mean 16	12.3	4.6	3.0	1.8	9.3	3.5	1.2			25
σ	2.7	0.38	0.40	0.29		0.66	0.31			

All measurements except row no. are in mm, not corrected for carbonization. Means and standard deviations of five measurements and two estimates for six Preclassic Cuello cupule categories. All the measurements (mm) are from carbonized specimens, so for comparison to uncarbonized specimens, about 25% needs to be added to the values.

* Row no. estimated: 360°/cupule angle, rounded off, multiplied by 2.

† Rachis diameter estimated by law of cosines: rachis diameter = cupule width × 1.414 $(1\text{-cosine cupule angle})^{-0.5}$.

Excavations at Cuello during 1976–1980 revealed superimposed early structures, some of probably ceremonial function[1-3]. The overall chronology of the site is from about 2000 b.c. in radiocarbon years to about AD700 in calendar years, the early part of the Late Classic period of historically dated Maya civilization[4,5]. (The Preclassic or Formative period before AD 250 is dated by a suite of 20 radiocarbon dates from the Cambridge (Q) and Los Angeles (UCLA) laboratories provided by V. R. Switsur and R. Berger; further samples are being processed by these two laboratories and also that at La Jolla (LJ) by H. Suess. (Lower case b.c. and a.d. indicate radiocarbon years with no correction for the bristlecone pine chronology, geomagnetic curve, or other factors.) The occupation of Cuello is formally divided into five successive phases, three in the Formative and two in the Classic period, based on the stratigraphy, pottery and radiocarbon dates from excavations principally in Platform 34 at the site. The Early Formative Swasey phase (2000–1000 b.c.) is succeeded by the Middle Formative Lopex Mamom phase (1000–400 b.c.) and the Late Formative Cocos Chicanel phase (400 b.c.–AD 250), then by the Early Classic Nuevo Tzakol (AD 250–600) and Late Classic Santana Tepeu (AD 600–700). Classic period deposits consist of superficial architecture and do not feature in the sampling programme reported here.

For the entire Formative period a succession of plaster floors define and separate layers of occupational material and fill. The possibility of vertical displacement of plant remains after deposition is minimal, although some animal and insect burrows were noted in the excavation and some root penetration was also present. The recycling of earlier material in later construction fills is, on the other hand, a reality attested to by abraded sherds of Swasey pottery occurring in Mamom and Chicanel contexts.

From 1978 to 1980 we carried out an intensive programme of sampling for plant remains by water flotation[6], processing 250 50-l samples representing more than 10 metric tons of earth. Soil samples were sealed into plastic bags and before flotation were partially air-dried to enhance recovery of carbonized plant parts, mollusc shells and bones. The smallest materials were land snails and seeds whose longest dimensions are 0.5 mm, so that little recoverable and identifiable material was missed. We recovered *Zea mays* kernel, cupule and glume fragments from all contexts including the earliest occupation on the old land surface dated at 1950±65 b.c. (Q-1571)[1]. From the samples analysed so far, 141 cupules (of a total of 329) and 56 kernels (of 821) were complete enough to measure in several ways. The lack of intact maize cobs may be due to breakage during construction of the platforms, predepositional handling including burning, and a low level of cob induration.

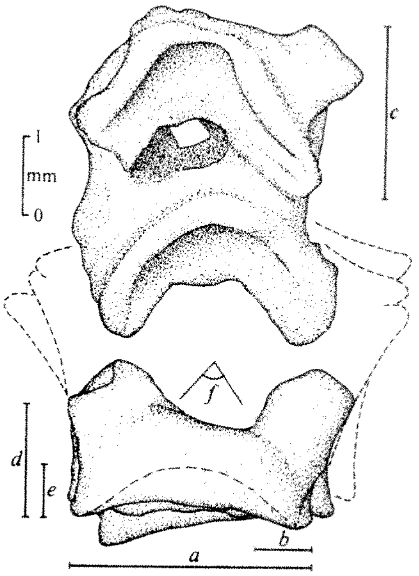


Fig. 1 A cob fragment with two almost complete cupules and one partial cupule to each side, similar to some Swasey type 2 cupules, dating to about 1000 b.c. Upper, side view; lower, end view (apical) demonstrating that the whole cob had five ranks of cupules, each subtending two fertile spikelets, giving ten rows of kernels. An estimate of the rachis outline and the interior of one cupule are indicated by dashes on the lower part of the fiugre. The measurements indicated are *a*, cupule width; *b*, cupule wing width; *c*, cupule length; *d*, cupule thickness; *e*, cupule depth, and *f*, cupule angle (used for estimating row number). (Drawn by R.McK.B.)

The variables we used to measure the cupules (Fig. 1) have been applied to a wide range of prehistoric and modern cobs[7–11]. Cupule angle has been used by Cutler to estimate row number[12]. Usually several measurements need to be used simultaneously to distinguish types of maize because races overlap in many metric values, often being differentiated by only one of the available variables. Several characters that are not easily quantifiable, such as cupule shape, may also aid in separation. Cuello maize types are named according to the phases in which they first appear.

The classification of extant races of maize is based on a broad spectrum of variables: morphology of the plant, tassel and ear; physiology; cytology; and most recently, isozymes of the coleoptile[13–16]. Because the Cuello collection consists of only charred kernel and cob fragments, much of the necessary evidence for complete racial classification is lacking. Nevertheless, it is important to attempt to understand the relationships of the Cuello maize to existing modern races and other archaeological collections.

The smallest type of Cuello maize is Swasey type 1 (SW-1). It has small elliptical to rectangular cupules with narrow cupule wings (Table 1), and 10–16 rows of grains, as estimated by cupule angle. The smallest kernels from Cuello, probably associated with this type, all have hard endosperm and probably represent a popcorn. They measure 4.2×3.2 mm (kernel width × kernel thickness, charred). Allowing for 20% shrinkage on carbonization, and estimating rachis diameter using the law of cosines (equation, Table 1), we believe that the intact ear would have been about 22 mm in diameter, about the size of a modern US popcorn called Ladyfinger. Without an entire cob, it is impossible to estimate length. SW-1 is more like early types of maize from South America (Cupisnique types 3 and 4, Gallinazo types 1 and 3) than any reported Mexican archaeological collections (Table 2).

Swasey type 2 (SW-2) and Swasey type 3 (SW-3) are somewhat similar, but can be distinguished by cupule width and shape. In outline SW-2 cupules are trapezoidal, chevron-shaped (Fig. 1) or B-shaped. SW-3 cupules tend to be more rectangular with wide wings, some of which resemble small 'ears'. The cupules of SW-2 often have slight apical protrusions resembling the toe of a boot. This configuration (sometimes called 'shoe-shaped' or 'Spanish stirrup') is present in some Tehuacán specimens, as well as modern Chapalote and many South American races. SW-2 has 10–16 rows of kernels averaging 5.6×4.2 mm (charred) which resemble a rounded popcorn. SW-3 has a slightly lower mean row number (with a range of 8–16) and kernels measuring 6.2×4.8 mm (charred) which seem to represent both popcorn and hard flint types. SW-2 is somewhat like Abejas maize but is very like Gallinazo types 1 and 4 from the north coast of Peru and Guadalupe-Largas maize from Oaxaca. SW-3 also resembles Guadalupe-Largas but is more like Early Nal Tel B and Pre-Chapalote (Table 2).

Mamom type 1 (MA-1) is larger than Swasey maize groups. It has wide, but comparatively short, rectangular to 'eared' cupules with broad wings. Charred kernels apparently associated with this type are all flinty (probably a soft flint), measuring 6.4×4.6 mm (charred). Uncarbonized MA-1 cobs would have been about 16 mm in diameter and, with kernels, the ears may have been about 28 mm in diameter. MA-1 is almost identical to Early Nal Tel B from La Perra Cave (Table 2) and is similar to Early Teosintoid, Late Nal Tel and Early Chapalote. With the appearance of MA-1, the resemblance between the Cuello material and the listed Peruvian maize types diminishes. MA-1 was probably very similar to a previously undescribed Central American race which we have collected from Santa Cruz, August Pine Ridge and Rancho Creek (Orange Walk District) and Aguacate (Toledo District) in Belize. This race has 10–14 rows of soft flint or semi-hard dent kernels with white, purple, yellow or pale red aleurone. The ears average 32 mm in diameter and 106 mm long. Although local Mayan farmers refer to this race as 'Dzit Bacal', it bears no resemblance to published descriptions of Dzit Bacal from Guatemala or the Yucatan[15,16].

Mamom type 2 (MA-2) is the largest of all Preclassic Cuello maize varieties. MA-2 has deep, wide trapezoidal or chevron-shaped cupules with 8–12 rows of flinty kernels measuring 6.8×4.3 mm (charred). It resembles Early Nal Tel B and Late

Table 2 Formative period maize from Latin America

Area	Type	Time range	Mean kernel row no.	Mean cupule width (mm)	Mean cupule length (mm)	Mean cupule wing width (mm)	Mean rachis diameter (mm)	Mean cob length (mm)	Index of similarity to Cuello maize types: SW-1	SW-2	SW-3	MA-1	MA-2	CH-1
Tehuacán Valley	Coxcatlán	5000–3500 b.c.	7.6	3.0	2.7	0.4	3.5	23.5	17	33	17	17	17	0
	Abejas	3500–2300 b.c.	9.0	4.2	2.5	0.4	5.1	27.7	0	50	33	17	33	0
	Early Teosintoid	1500–900 b.c.	12.0	6.2	3.5	1.3	12.4	78.4	33	33	33	83	83	83
Sierra de Tamaulipas	Early Nal Tel A	2495–250 b.c.	8.9	5.4			8.4	48.0	0	33	33	33	33	33
	Early Nal Tel B	2495–250 b.c.	10.2	6.2	3.1	1.7	10.6	60.0	20	20	80	100	80	80
	Late Nal Tel	250 b.c.–a.d.150	10.9	7.1	2.6	1.6	13.2	86.0	20	20	60	80	80	40
Infiernillo Canyon	Pre-Chapalote	2350–400 b.c.	9.0	4.8*	3.0	1.4	7.5	30.0	20	20	80	60	40	40
	Early Chapalote	1850 b.c.–a.d. 800	10.5	5.8*	4.1	1.8	9.9	58.4	20	20	60	80	20	100
Oaxaca Valley	Guadalupe-Largas	1200–550 b.c.	11.5	(3.3)	(3.2)		(6.6)*		25	75	75	25	0	50
Chicama Valley[8]	Cupisnique-3	850–450 b.c.	15.3	2.6	1.7	0.5	7.3	49.5	80	40	20	0	0	0
	Cupisnique-4	850–450 b.c.	12.0	2.8	1.6	0.5	6.0	46.0	80	20	20	20	20	20
	Gallinazo-1	a.d. 20–220	15.1	2.9	1.8	0.5	7.6	45.0	80	80	20	0	0	0
	Gallinazo-2	a.d. 20–220	13.7	4.0	1.7	0.6	8.6	46.0	40	60	60	40	0	20
	Gallinazo-3	a.d. 20–220	10.2	2.9	1.7	0.5	5.3	36.0	80	20	20	20	20	0
	Gallinazo-4	a.d. 20–220	10.0	4.0	2.1	0.8	6.7	42.0	60	80	60	20	20	0
	Gallinazo-5	a.d. 20–220	11.1	5.4	1.7	1.0	9.5	62.0	60	60	40	60	20	40

Measurement means for early maize types from eight archaeological sites in five areas of Latin America. The data from Tamaulipas are taken from photographs and published information. The Tehuacán Valley (Puebla, Mexico) sites are the Coxcatlán and San Marcos Caves[8]; the Sierra de Tamaulipas (Mexico) site is La Perra Cave[17]; the Infiernillo Canyon (Mexico) sites are Romero's Cave, Valenzuela's Cave and Ojo de Agua[18]; the Oaxaca Valley (Mexico) site is Fabrica San José[19]; the Chicama Valley (Peruvian north coast) site is Huaca Prieta, including an associated area 150 m north[8]. The maize comes from two or more cultural phases at each site. An asterisk indicates that the value for rachis diameter is estimated as explained in Table 1, or that cupule width is estimated from rachis diameter. Figures in parentheses are from carbonized specimens with about 20% shrinkage. Cupule depth for the five Coxcatlán cobs is 0.7 mm, and for the four Abejas cobs is 1.1 mm. Cupule thickness for Coxcatlán maize averages 1.7 mm. The index of similarity is calculated by adding 25% to the values for carbonized Cuello specimens, then establishing a ±20% range around these adjusted values, counting overlaps of this range with the mean values per type, and dividing by the number of possible comparisons.

Nal Tel from La Perra Cava and Early Teosintoid maize from Tehuacán (Table 2). Before carbonization, the cobs of MA-2 would have been about 20 mm in diameter and with kernels the ears would have measured about 30 mm in diameter. It was probably quite similar to the Tuxpeño-introgressed Nal Tel grown by traditional *milperos* throughout Belize today.

Approximately three-quarters of the Cuello Chicanel-1 material comes from Square 45/35 Layer 1597, a burnt house (apparently with a corn crib) dating to the early Chicanel phase, about 250 b.c. Because all the maize from this context was quite similar, it probably represents a single type from one harvest. CH-1 has open, trapezoidal, 'Spanish-stirrup' cupules with broad cupule wings and 8–16 rows of flinty kernels measuring 6.5 × 5.0 mm (charred). It is almost identical to Early Chapalote from Tamaulipas and is quite similar to Early Teosintoid and Early Nal Tel B (Table 2).

Because they lack all but cob and kernel fragments, we know nothing about ear length, plant form, productivity or other characters of these early lowland maize types. Although the general increase in rachis diameter and kernel size through time suggests an increase in productivity, we must be cautious in making this assumption. A single plant with many small ears (which is true of many popcorns and primitive maize types) could produce as much food as a plant with one or two larger ears. Perhaps evidence for other plant characters and the early adaptation of maize to the lowlands will be found in dry caves in arid valleys along the upper Motagua drainage in southeastern Guatemala.

The earliest domesticated maize was probably adapted to the semi-arid highlands of central and southern Mexico. By at least 2000 b.c. a distinct lineage including SW-1 and SW-2 was developed which was more adapted to the humid lowlands. This lineage probably spread throughout tropical Central America and into coastal South America and contributed to the development of sedentary horticulturalists such as the Swasey people at Cuello. After 1000 b.c. the relationship of Cuello maize to highland Mexican types is more pronounced. The improvement of earlier Swasey maize and hybridization with highland germplasm (and possibly teosinte) produced new varieties which eventually helped to support a dense population and the complex culture of the Maya lowlands.

Major funding for this field work came from the National Geographic Society, the British Museum, the Society of Antiquaries of London, the Wenner-Gren Foundation for Anthropological Research Inc. and Rutgers University. Valuable field assistance was given by Ingrid Wuebber, Kathryn Elsesser, and other members of the Cuello Project Staff; John Masson and Richard Wilk assisted with ethnobotanical collection of modern Maya maize. We thank the Director of the Museo Regional de Puebla and Mario Camacho Robles, Curador General, for access to collections of archaeological maize from the Tehuacán Valley, Puebla, Mexico.

Received 4 February, accepted 31 October 1980.

1 Hammond, N (ed) *Cuello Project 1978 Interim Report* Publ no. 1 (Archaeological Research Program, Douglass College, Rutgers University, 1979)
2 Hammond, N *Belizean Stud* **8**, 33–44 (1980), *Belizean Stud* (in the press)
3 Hammond, N *et al Am Antiq* **44**, 92–110 (1979)
4 Hammond, N *et al Nature* **260**, 579–581 (1976), **267**, 608–610 (1977)
5 Hammond, N *Am Antiq* **44**, 92–110 (1979)
6 Minnis, P E & LeBlanc, S *Am Antiq* **41**, 491–493 (1976)
7 Bird, R McK. thesis, Univ California, microfilm 71–9767 (1970)
8 Bird, R McK *Maize Genet Coop News Lett* **52**, 90–92 (1978)
9 Bird, R McK *Maize Genet Coop News Lett* **53**, 53–54 (1979)
10 Bird, R McK & Bird, J B *Am Antiq* **45**, 325–332 (1980)
11 Nickerson, N H *Ann Mo Bot Garden* **40**, 79–111 (1953)
12 Cutler, H. C *Univ Utah Anthr Pap* **80**, 10–11 (1966)
13 Bird, R. McK & Goodman, M M *Econ Bot* **31**, 471–481 (1978)
14 Smith, J S C & Lester, R N *Econ Bot* **34**, 201–218 (1980)
15 Wellhausen, E J *et al Races of Maize in Mexico* (Bussey Institute, Harvard University, Cambridge, 1952)
16 Wellhausen, E J *Races of Maize in Central America*, Publ 511 (National Academy of Sciences—National Research Council, Washington DC, 1957)
17 Mangelsdorf, P C *et al Bot Mus Leafl, Harvard Univ* **17**, 125–150 (1956).
18 Mangelsdorf, P C *et al Bot Mus Leafl, Harvard Univ* **22**, 33–62 (1967)
19 Ford, R I in *Prehistory and Human Ecology of the Valley of Oaxaca* (ed Flannery, K V) Mem. Mus Anthr Univ Mich no 8, 4, 261–268 (1976)

Electrical activity following cellular recognition of self and non-self in a sea anemone

Roger Lubbock

Department of Zoology, Downing Street, Cambridge CB2 3EJ, UK

G. A. B. Shelton

Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

The sea anemone *Anthopleura elegantissima* lives in clonal colonies[1–3] and possesses a cellular recognition system of remarkable specificity[4], by which it can recognize members of its own clone; other anemones, including individuals of the same species which are not syngeneic, are attacked. Attack is initiated by contact with a foreign anthozoan and involves the inflation of specialized tentacle-like structures known as acrorhagi, which contain numerous stinging cells[1,2]. These stinging cells only discharge when the tip of the acrorhagus is in physical contact with the surface of a foreign anthozoan; contact with syngeneic individuals, organisms other than anthozoans and inanimate objects does not elicit discharge[4]. We show here that the recognition of allogeneic tissue is accompanied by a novel form of local electrical activity in the acrorhagus that is usually, but not invariably, followed by nematocyst discharge. This type of electrical activity was not found during contact with syngeneic tissue or inanimate objects and seemed to be a consequence of the recognition of allogeneic surface markers by cells at the tip of the acrorhagus.

Clones of *A. elegantissima* were maintained in several 150–400 l aquaria containing artificial sea water (Tropic Marin Neu; relative density 1.025) at 23 °C. Before experimentation, an anemone was placed in a small bowl of sea water and induced to inflate its acrorhagi by contact with an allogeneic individual. Extracellular suction electrodes[5,6] (tip diameter 200 μm) were applied to a point about half way along the inflated acrorhagus; signals were amplified by Tektronix Type FM 122 differential preamplifiers and displayed on a Tektronix Type 5111 storage oscilloscope and a Washington curvilinear pen recorder. Care was taken to maintain the electrodes in a single position when comparing responses. Discharge of nematocysts was observed using a binocular microscope, usually with dark ground illumination.

Touching the acrorhagus with the intact surface of excised syngeneic tentacles consistently failed to induce either nematocyst discharge or any recordable electrical excitation; in addition, no detectable activity could be produced by touching the acrorhagus with inanimate objects (metal forceps or glass rod). On the other hand, contact of the acrorhagial tip with an allogeneic tentacle required only a short delay before it produced massive, simultaneous nematocyst discharge, including numerous nematocysts remote from the point of contact (Fig. 1*a*, *b*); discharge was accompanied by a triphasic spike up to 50 μV in amplitude and about 350 ms in duration (Fig. 2). The precise timing of this spike in relation to that of nematocyst discharge was not established and it was therefore unclear whether the spike initiated discharge or was a consequence of it. Of greater interest in the present context was the occurrence of a variable number of much smaller impulses of up to 10 μV and up to 50 ms duration, which preceded the spike. It was particularly notable that these small impulses could also be evoked by brief contact with allogeneic material, even though no discharge of acrorhagial nematocysts and no spike occurred (Fig. 2). These impulses could not be ascribed to any of the known sea anemone electrical conduction systems[7]; they were distinctive in that they seemed to be confined to a single acrorhagus and could be

0028-0836/81/020059—02$01.00

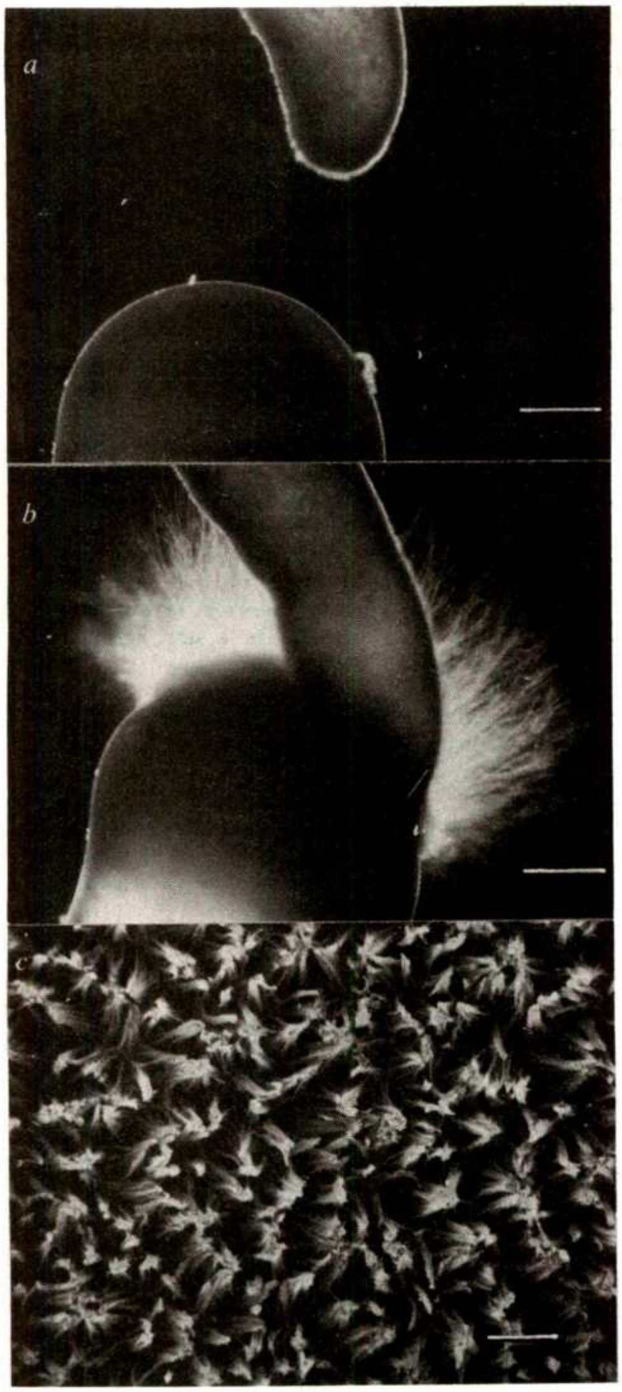


Fig. 1 *a*, Acrorhagus (below) and allogeneic tentacle (above) before contact. Scale bar, 300 μm. *b*, Massive, simultaneous discharge of acrorhagial nematocysts after contact between acrorhagus (below) and allogeneic tentacle (above); discharge usually occurred within 0.5–3 s of contact. Scale bar, 300 μm. *c*, Scanning electron micrograph of ciliary cones on the tip of an inflated acrorhagus. Scale bar, 5 μm.

shown, using two electrodes, to decrease rapidly in amplitude at increasing distance from the acrorhagial tip. They were clearly not the result of the foreign tentacle stinging the acrorhagus: the impulses could also be obtained using fresh pieces of allogeneic column tissue (largely devoid of nematocysts), as well as washed foreign tentacles previously fixed in formalin. It thus seemed that the small impulses were a consequence of cells at the acrorhagial tip recognizing allogeneic markers on the foreign tissue.

The receptors associated with recognition were presumed to reside on the ciliary cones that densely cover the distal portion of the acrorhagus (Fig. 1*c*); it was notable that touching the unciliated sides and base of the acrorhagus with allogeneic tissue failed to produce any detectable electrical activity.

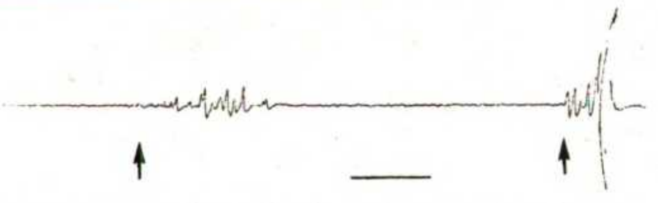


Fig. 2 Pen recording showing electrical response of acrorhagus to contact with an allogeneic tentacle. The beginning of each period of contact is indicated by an arrow; duration of contact was ~1 s. The first period of contact (left-hand arrow) evoked small impulses but no nematocyst discharge. The second period of contact (right-hand arrow) evoked small impulses which were succeeded by a large triphasic spike; the spike was accompanied by massive nematocyst discharge. Scale bar, 1 s.

We suggest the following hypothesis of the means by which recognition of a foreign anemone leads to nematocyst discharge: specific molecules on the surface of the allogeneic tissue bind to surface receptors on cells at the tip of the acrorhagus, inducing electrical excitation; local events of this kind summate until a threshold is reached, at which point massive nematocyst discharge occurs. The fact that electrical activity was recorded on contact with allogeneic tissue, but not with syngeneic tissue of inanimate objects, is consistent with the previously advanced theory[8] that self recognition in anemones is an essentially passive process, with specific receptors failing to be excited, whereas non-self recognition is an active process involving receptor excitation.

We thank W. B. Amos and G. Koch for commenting on the manuscript.

Received 18 July; accepted 24 October 1980.

1. Francis, L. *Biol. Bull.* **144,** 64–72 (1973).
2. Francis, L. *Biol. Bull.* **144,** 73–92 (1973).
3. Francis, L. *Am. Zool.* **19,** 669–681 (1979).
4. Lubbock, R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 6667–6669 (1980).
5. Josephson, R. K. *J. exp. Biol.* **45,** 305–320 (1966).
6. Macfarlane, I. D. *J. exp. Biol.* **51,** 377–385 (1969).
7. Shelton, G. A. B. in *Electrical Conduction and Behaviour in "Simple" Invertebrates* (ed. Shelton, G. A. B.) (Oxford University Press, 1980).
8. Lubbock, R. *J. exp. Biol.* **83,** 283–292 (1979).

Surface antigen differentiation during human myogenesis in culture

Frank S. Walsh

Muscular Dystrophy Research Laboratories, Institute of Neurology,
Queen Square, London WC1N 3BG, UK

Mary A. Ritter

Membrane Immunology Laboratory, Imperial Cancer Research
Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

The fusion of mononucleate myoblast cells into multinucleate myotubes (myogenesis) has often been studied as a model system of terminal cellular differentiation[1]. Although cell-surface changes during myogenesis have attracted much attention and a variety of surface antigens including the nicotinic cholinergic receptor[2], acetylcholinesterase[3] and muscle lectins[4,5] have been shown to be present on muscle cells, a detailed identification of markers specific to the various cell types has not been attempted. This is mainly because fibroblasts are a major contaminating cell type in primary muscle cultures and these cells have no known distinctive cell-surface antigen(s). For instance, surface fibronectin has been used to distinguish fibroblasts in the rat peripheral nervous system *in vitro*[6] but has proved ineffective in the human muscle cell system[7]. In addition, Lesley and Lennon[8] found that Thy-1 antigen was present on myoblasts but not myotubes of the rat L6 muscle cell line and primary cultures. However, Thy-1 antigen is also present on

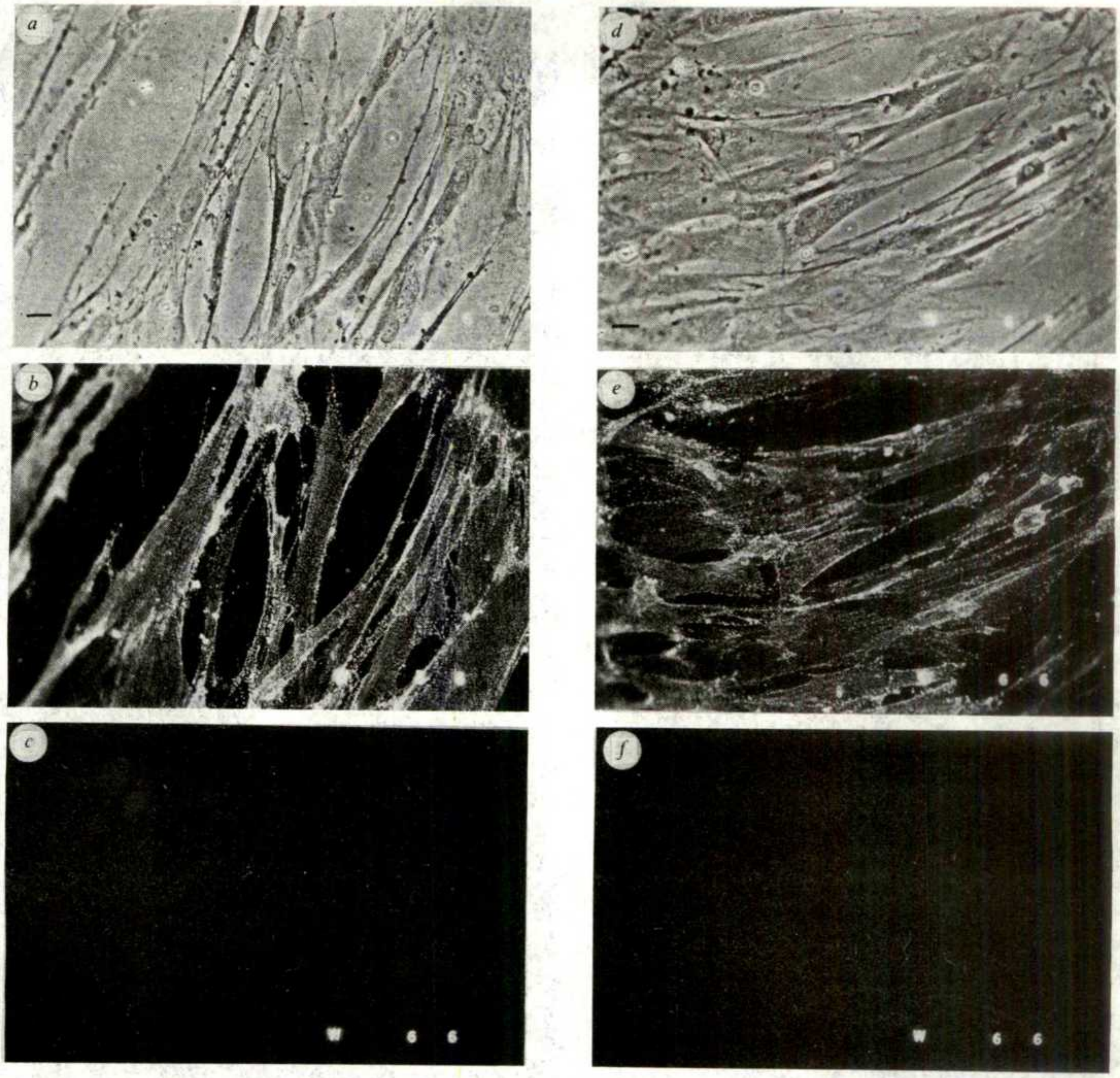
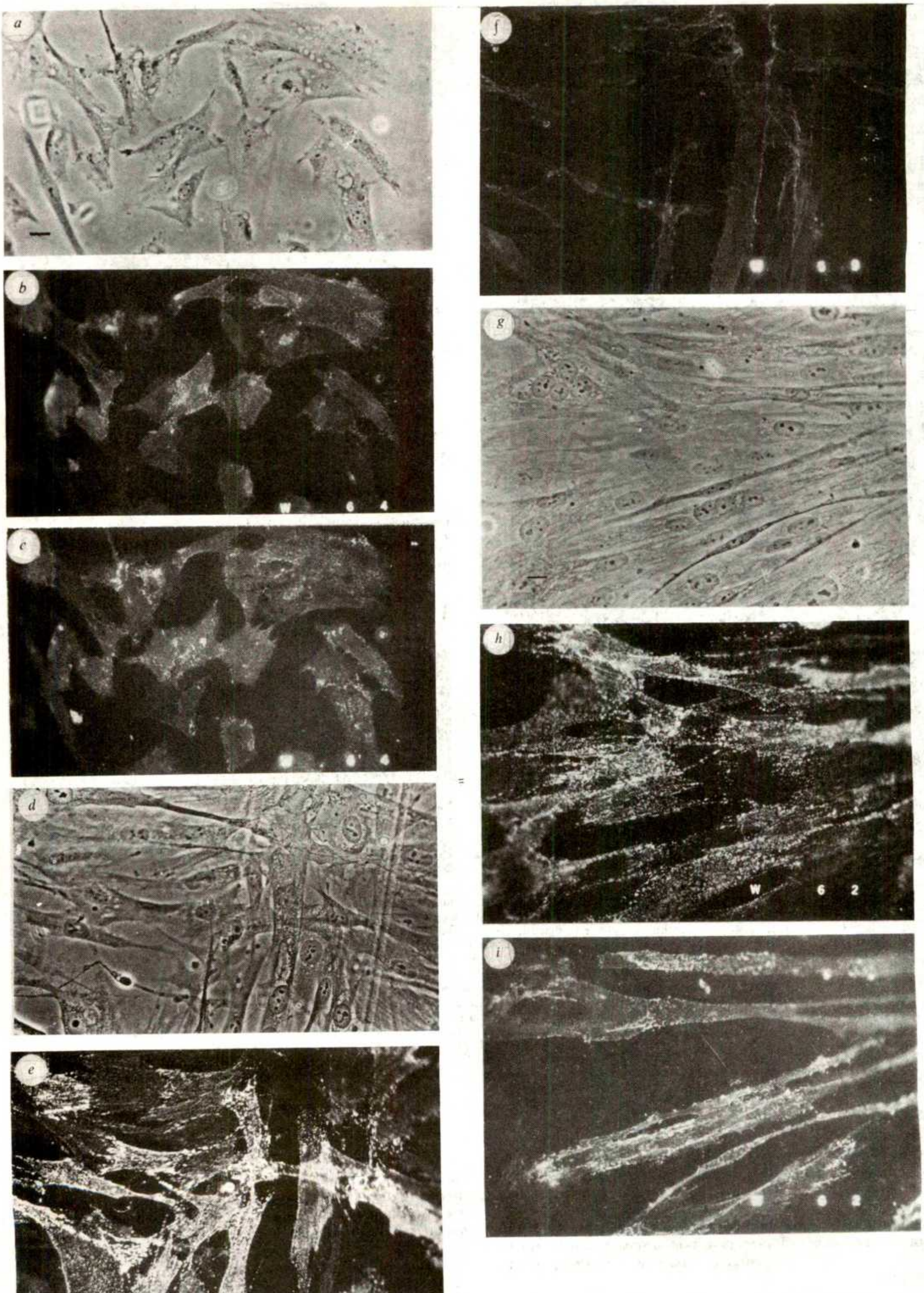


Fig. 1 Double immunofluorescent staining of fetal human skin fibroblast or muscle fibroblast cells with rabbit anti-rat Thy-1 antibody and 5.1 H11 monoclonal antibody. Fibroblasts were prepared from fetal skin by dissociating with trypsin and collagenase, using the procedure previously described for muscle[17]. The cells were grown on 60-mm Petriperm plates (Heraeus) in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS). Fetal human muscle fibroblasts were prepared by preplating[14] a single-cell suspension of enzymatically dissociated skeletal muscle. The cells that adhered to uncoated 100-mm Lux Petri dishes after 30 min incubation were regarded as muscle fibroblasts. These cells were grown to confluence in DMEM, 10% FCS, at least twice before use to allow any residual myoblasts to fuse into myotubes. These cultures express negligible activity of the muscle enzymatic marker creatine phosphokinase when grown to confluence. Immunofluoroescent staining of skin or muscle fibroblast cultures was carried out on subconfluent cultures. The Petriperm plates were cut into 10-mm^2 pieces and 10 μl of an F(ab')$_2$ preparation of rabbit anti-rat Thy-1 antibody (24 μg ml^{-1}) in phosphate-buffered saline (PBS, 10 mM Na phosphate buffer, 0.15 M NaCl, *p*H 7.3) and 10 μl of 1 : 100 dilution of 5.1 H11 antibody (ascitic fluid) in PBS added for 45 min at 20 °C. The cells were washed in PBS and 10 μl of an F(ab')$_2$ fragment of sheep anti-rabbit IgG antibody (60 μg ml^{-1}) labelled with rhodamine plus 10 μl of a 1 : 10 dilution of goat anti-mouse IgG antibody labelled with fluorescein (Meloy) added for 45 min at 20 °C. The cells were washed extensively in PBS, fixed in 3.5% formaldehyde in PBS for 5 min followed by 5% acetic acid in ethanol for 5 min at 20 °C. The samples were mounted on glass slides in glycerol and sealed with nail varnish. Samples were examined with a Leitz dialux 20 microscope equipped with phase contrast fluorescein and rhodamine optics and epi-illumination. *a*, Phase-contrast view of skin fibroblast culture; *b*, positive Thy-1 staining viewed with rhodamine optics; *c*, negative 5.1 H11 staining with fluorescein optics; *d*, phase-contrast view of muscle fibroblast culture; *e*, positive Thy-1 staining viewed with rhodamine optics; *f*, negative 5.1 H11 staining with fluorescein optics. Scale bar, 20 μm.

rat fibroblasts[8,9]. Thus, unequivocal identification of the major mononucleate cell types in muscle cultures, fibroblasts and myoblasts has not been possible. We report here the use of two surface antigens, Thy-1 and a muscle-specific antigen recognized by a monoclonal antibody, to identify unambiguously the four major types of cells present in human muscle cultures and to propose a model of cell-surface differentiation during human myogenesis.

Two antibodies were used in this study. The first is a rabbit antibody to purified rat brain Thy-1 (refs 10, 11) which we have used to detect the human homologue of Thy-1 (refs 12, 13). Preabsorption of this antibody with pure rat brain Thy-1 (1:3 molar ratio) abolished all immunofluorescence on human muscle cells (see below), showing the staining to be specific for the Thy-1 molecule. In addition, because previous absorption of the anti-Thy-1 with Thy-1.2 mouse thymocytes also abrogated


a
b
W 6 4
c
d
e
f
g
h
W 6 2
i

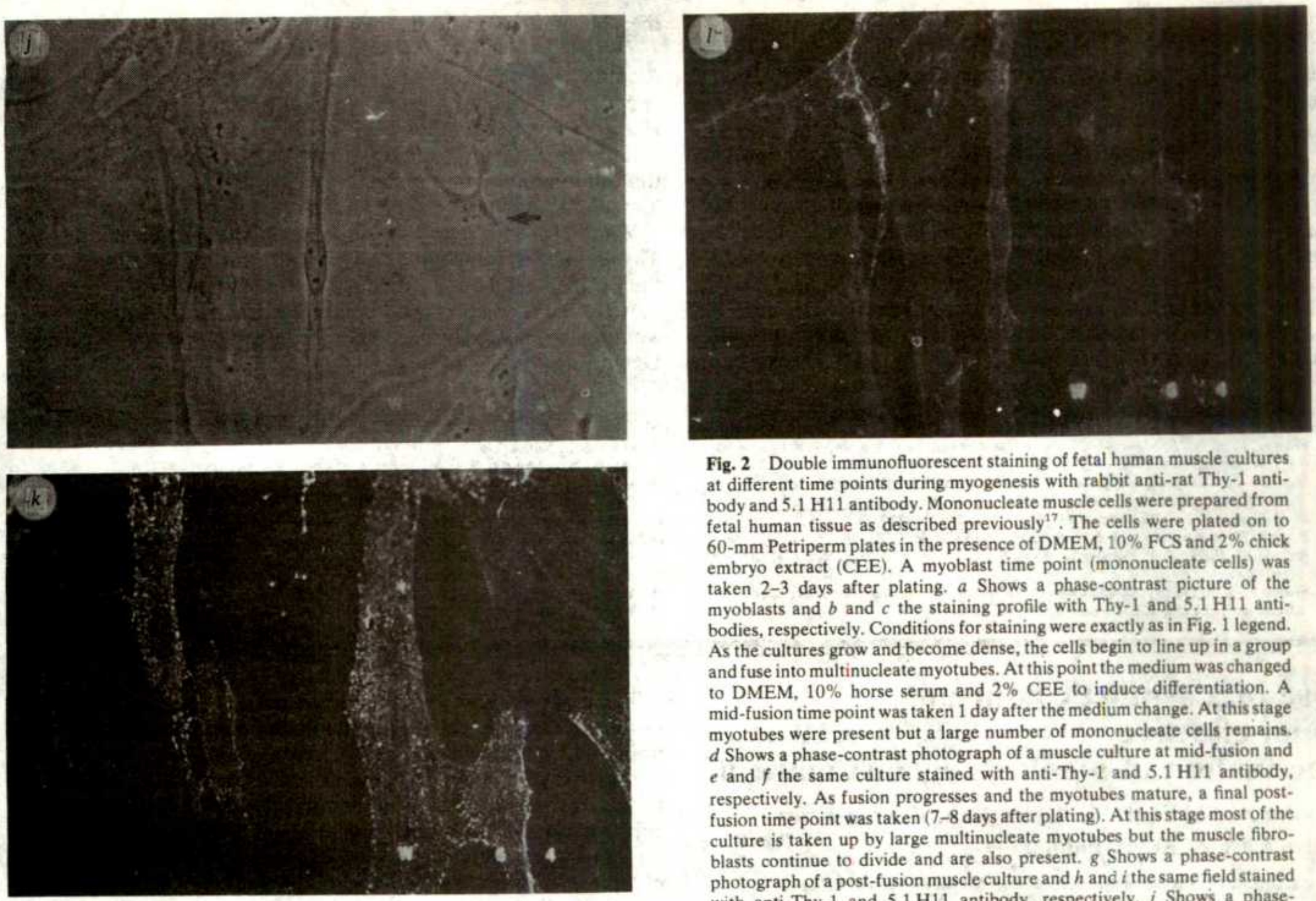


Fig. 2 Double immunofluorescent staining of fetal human muscle cultures at different time points during myogenesis with rabbit anti-rat Thy-1 antibody and 5.1 H11 antibody. Mononucleate muscle cells were prepared from fetal human tissue as described previously[17]. The cells were plated on to 60-mm Petriperm plates in the presence of DMEM, 10% FCS and 2% chick embryo extract (CEE). A myoblast time point (mononucleate cells) was taken 2–3 days after plating. *a* Shows a phase-contrast picture of the myoblasts and *b* and *c* the staining profile with Thy-1 and 5.1 H11 antibodies, respectively. Conditions for staining were exactly as in Fig. 1 legend. As the cultures grow and become dense, the cells begin to line up in a group and fuse into multinucleate myotubes. At this point the medium was changed to DMEM, 10% horse serum and 2% CEE to induce differentiation. A mid-fusion time point was taken 1 day after the medium change. At this stage myotubes were present but a large number of mononucleate cells remains. *d* Shows a phase-contrast photograph of a muscle culture at mid-fusion and *e* and *f* the same culture stained with anti-Thy-1 and 5.1 H11 antibody, respectively. As fusion progresses and the myotubes mature, a final post-fusion time point was taken (7–8 days after plating). At this stage most of the culture is taken up by large multinucleate myotubes but the muscle fibroblasts continue to divide and are also present. *g* Shows a phase-contrast photograph of a post-fusion muscle culture and *h* and *i* the same field stained with anti-Thy-1 and 5.1 H11 antibody, respectively. *j* Shows a phase-contrast photograph of a mid-fusion culture, and *k* and *l* the same field stained with anti-Thy-1 and 5.1 H11 antibodies, respectively. Note the mononucleate cell (see arrow) stained by 5.1 H11 antibody but not anti-Thy-1. Scale bar, 20 μm.

all immunofluorescent staining on human cells, it must be the rat–mouse cross-reacting antigenic determinant that is shared by rat and human Thy-1. The second reagent, 5.1 H11 monoclonal antibody[7], was produced from a cell fusion involving P3X63 Ag8 mouse myeloma cells and spleen cells from a BALB/c mouse immunized with human myotube cells. It reacts with human myoblasts and myotubes but not with fibroblasts and is human muscle cell specific[7].

Human fetal skin fibroblasts, muscle fibroblasts and muscle cultures at different stages of myogenesis were grown and tested for reactivity with these antibodies by indirect immunofluorescence. Figure 1 shows that Thy-1 antigen is present on all cells in human skin and muscle fibroblast cultures; neither of these cell types showed detectable reactivity with 5.1 H11 antibody. The results with anti-Thy-1 antibody agree with data from rats that show Thy-1 antigen is expressed on fibroblasts[9].

The distribution of Thy-1 and 5.1 H11 antigens in cultures of mononucleate muscle cells is shown in Fig. 2*a–c*. In contrast to the results with skin or muscle fibroblasts, the majority of the cells in the myoblast culture are labelled by both anti-Thy-1 and 5.1 H11 antibodies. A few (10%) of the cells were found to be Thy-1$^+$/5.1 H11$^-$. These muscle cultures have been significantly depleted of fibroblasts by preplating[14]; we therefore believe that the Thy-1$^+$/5.1 H11$^-$ cells are muscle fibroblasts, because these cells express the same antigenic phenotype as the fibroblast cultures. At the mid-fusion stage of growth and differentiation (Fig. 2*d*), the mononucleate myoblast cells have started to fuse into multinucleate myotubes, although some mononucleate cells remain. All the multinucleate myotubes in the cultures are stained by 5.1 H11 antibody (Fig. 2*f*) but not by anti-Thy-1 antibodies (Fig. 2*e*). Multinucleate myotubes are therefore Thy-1$^-$/5.1 H11$^+$. As the myoblast population has mostly fused to form myotubes, there are few myoblasts of Thy-1$^+$/5.1 H11$^+$ phenotype at this time. Most of the mononucleate cells in these cultures are Thy-1$^+$/5.1 H11$^-$ and, we believe, fibroblasts. At a later stage in muscle differentiation, post-fusion, the fibroblast population has enlarged and the majority of the mononucleate cells are Thy-1$^+$/5.1 H11$^-$. The myotubes here are still Thy-1$^-$/5.1 H11$^+$. In general, the strength of 5.1 H11 staining is much higher in post-fusion than in mid-fusion cultures so it is possible that the level of the antigen increases during differentiation. Two further points deserve comment. First, in mid-fusion and post-fusion cultures a small percentage of mononucleate cells is often found expressing the phenotype Thy-1$^-$/5.1 H11$^+$. Figure 2*j, k, l* shows an example of this behaviour and it seems likely that these cells are post-mitotic myoblasts which do not fuse to myotubes[15].

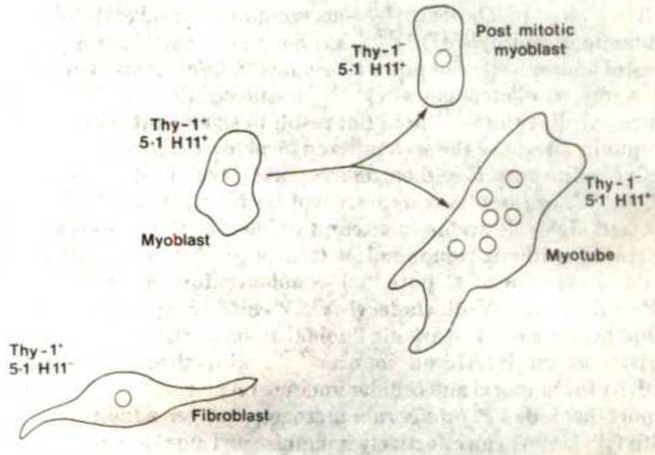


Fig. 3 A scheme for cell-surface differentiation during human myogenesis.

Second, we have confirmed the conclusions presented for the primary muscle cultures in experiments carried out on myogenic single-cell clones derived from human skeletal muscle.[16] Thus, the pure populations of myoblasts in these clones are all Thy-1$^+$/5.1 H11$^+$ and multinucleate myotubes are Thy-1$^-$/5.1 H11$^+$. No Thy-1$^+$/5.1 H11$^-$ cells were found in either pre- or post-fusion cultures.

Thus, we have found a distinctive phenotype for each of the main cellular populations in human skeletal muscle cultures. It is therefore possible to construct a model of surface membrane changes during myogenesis incorporating these results (Fig. 3). 5.1 H11 is present throughout myogenesis, whereas Thy-1 is lost from the cell surface as myoblasts fuse to form multinucleate myotubes (a similar situation is found with rat Thy-1[8]) or when they become post-mitotic myoblasts. Fibroblasts remain Thy-1$^+$/5.1 H11$^-$ throughout culture. Hence, it should be possible to separate out the various muscle cell types for biochemical analysis. A number of interesting points remain to be resolved. Exactly when is Thy-1 lost: before or after myotube formation? Could cell-surface Thy-1 be involved in the fusion process and is this differentiation pattern altered in diseases such as Duchenne muscular dystrophy? These aspects are now being investigated.

We thank Dr Roger Morris for the gift of the rabbit anti-rat Thy-1, rhodamine-conjugated sheep anti-rabbit antisera and the purified rat Thy-1 antigen, Dr Rose Yasin for the human muscle clones and Susheela Dhut for technical assistance. This work was supported by a special project grant from the Muscular Dystrophy Group of Great Britain to F.S.W.

Received 2 July; accepted 10 November 1980.

1. Pearson, M. L. in *The Molecular Genetics of Development* (eds Leighton, T. & Loomis, W. F.) (Academic, New York, in the press).
2. Merlie, J. P., Changeux, J. P. & Gros, F. *J. biol. Chem.* **253,** 2882–2891 (1978).
3. Shainberg, A. & Brik, H. *FEBS Lett.* **88,** 327–331 (1978).
4. Mir-Lechaire, F. J. & Barondes, S. H. *Nature* **272,** 256–258 (1978).
5. Podleski, T. R., Greenberg, I. & Nichols, S. C. *Expl Cell Res.* **122,** 305–316 (1979).
6. Raff, M. C., Fields, K. L., Mirsky, R., Pruss, R. M. & Winter, J. *Brain Res.* **174,** 283–308 (1979).
7. Walsh, F. S. in *Synaptic Constituents in Health and Disease* (eds Brzin, M., Sket, D. & Bachelard, H.) 285–320 (Pergamon, Oxford, 1980).
8. Lesley, J. F. & Lennon, V. A. *Nature* **268,** 163–165 (1977).
9. Stern, P. L. *Nature* **246,** 76–78 (1973).
10. Morris, R. J., Mancini, P. E. & Pfeiffer, S. E. *Brain Res.* **182,** 119–135 (1980).
11. Morris, R. J. & Ritter, M. A. *Cell Tissue Res.* **206,** 459–475 (1980).
12. Dalchau, R. & Fabre, J. W. *J. exp. Med.* **149,** 576–591 (1979).
13. Arndt, R., Stark, R., Klein, P. & Thiele, H.-G. *Immunology* **35,** 95–103 (1978).
14. Yaffe, D. in *Tissue Culture Methods and Application* (eds Kruse, P. F. & Patterson, M. K. Jr) 106–114 (Academic, New York, 1973).
15. Linkhart, T. A. & Hauschka, S. D. *Devl Biol.* **69,** 529–548 (1979).
16. Yasin, R., Kundu, D. & Thompson, E. J. *Cell Biol. int. Rep.* **4,** 783 (1980).
17. Yasin, R. *et al. J. neurol. Sci.* **32,** 347–360 (1977).

Use of a synthetic adjuvant in an effective vaccination of monkeys against malaria

Wasim A. Siddiqui, Siu-Chow Kan, Kenton Kramer, Steve Case & Kevin Palmer

Department of Tropical Medicine and Medical Microbiology,
John A. Burns School of Medicine, University of Hawaii,
Honolulu, Hawaii 96816

John F. Niblack

Department of Pharmacology, Central Research Division,
Pfizer Inc., Groton, Connecticut 06340

Research towards the development of a vaccine against malaria has recently accelerated[1], mainly because of a resurgence in the worldwide incidence of the disease[2]. These investigations have been aided by the use of the owl monkey for direct experimentation with human malaria[3–5], techniques for producing *Plasmodium falciparum* antigen through continuous or short-term *in-vitro* cultivation[6,7], and the recent demonstration of effective immunization of owl monkeys against *P. falciparum*[8,9]. However, the inclusion of a strong adjuvant in the vaccines seems to be required for the development of protective immunity. Thus, inclusion of the whole mycobacterial cell wall/mineral oil adjuvant (Freund's complete adjuvant: FCA)[8,9] or a semi-synthetic derivative of the active component of that cell wall (6-*O*-stearoyl-*N*-acetylmuramyl-L-alanyl-D-iso-glutamine: stearoyl-MDP)[10,11] has been necessary for the successful immunization of primates against *P. falciparum*. The use of X-ray irradiated parasites[12–14] or substitution of FCA by muramyl-dipeptide[15,16] does not result in significant protective immunity. Because the systemic and local toxic responses to the active adjuvants is either unacceptable (as in the case of FCA)[10,11], or essentially undescribed (as for stearoyl-MDP), it seemed highly desirable to attempt to identify alternative, well tolerated synthetic compounds with strong adjuvant activity for malarial vaccines. A potential candidate for this role was CP-20,961 (*N, N*-dioctadecyl-*N′*, *N′*-bis (2-hydroxyethyl)-propanediamine), a synthetic lipoidal amine, originally characterized as an interferon inducer[17–20], with strong adjuvant activity for humoral and cellular immune responses[21,22]. We now report that killed *P. falciparum* merozoite antigen administered with CP-20,961 can effectively immunize owl monkeys against a lethal *P. falciparum* infection.

The adjuvant action of CP-20,961 derives from an enhancement of leukocyte traffic through antigen deposits and into the draining lymph nodes, potentiating the expansion of immunoreactive clones[23]. Immunizations of guinea pigs[21], mice[23], rats, cows and horses (unpublished data) with intramuscular vaccines adjuvanted with CP-20,961 in aqueous soybean oil emulsions result in elevated immune responses, with only mild, transient swelling at the site of injection. The nature and extent of the cellular reactions induced at the site of CP-20,961 injection and in the draining lymphatics has been well characterized and seems to be quite benign, in contrast to reactions noted with equipotent doses of FCA[23]. Systemic toxicity of CP-20,961 also seems quite low when assessed either acutely ($LD_{50} > 2{,}000$ mg per kg, mice, intraperitoneally i.p.), or chronically (unpublished data). It was these characteristics which prompted our investigation of CP-20,961 as a substitute for FCA in an experimental malaria vaccine.

The Uganda-Palo Alto strain (FUP) of *P. falciparum* has been maintained in our laboratory by serial passages of blood-induced infections in owl monkeys (*Aotus trivirgatus griseimembra*)[24]. The immunizing antigen (mature segmenters containing individual merozoites) was obtained through short-term *in vitro* cultivation of owl monkey-derived *P. falciparum* parasites[11,25], and was concentrated and collected relatively free of other cellular elements[26]. The antigen preparation was stored at −20 °C for 2 months.

CP-20,961 was prepared with carrier soybean oil emulsion (Intralipid, Cutter Laboratories) as previously described[21]. The adjuvant stearoyl-MDP was used with carrier liposomes[10,27] and the method of malaria vaccine preparation using this adjuvant has been described earlier[11], as have the techniques for preparing malaria vaccine with FCA[8].

Nine owl monkeys were used, indicated by physical characteristics and fur coloration[28] to be of phenotype B. Using the standard chromosome preparation method[29] and well established classification criteria[28], six monkeys (A299, A304, A300, A297, A310 and A321) were found to belong to karyotype II and the remaining three (A320, A322 & A301) to karyotype III. Monkeys of both these karyotypes, originating in Columbia, South America, were found to be highly susceptible to blood challenge from a patent *Aotus* monkey infection of the FUP strain of *P. falciparum*[30].

Table 1 summarizes the composition of the vaccine and the vaccination schedule. Three monkeys were used as controls: A299 and A305 received parasite antigen only and A300 CP-20,961 adjuvant only. Of the six vaccinated monkeys,

0028-0836/81/020064—03$01.00

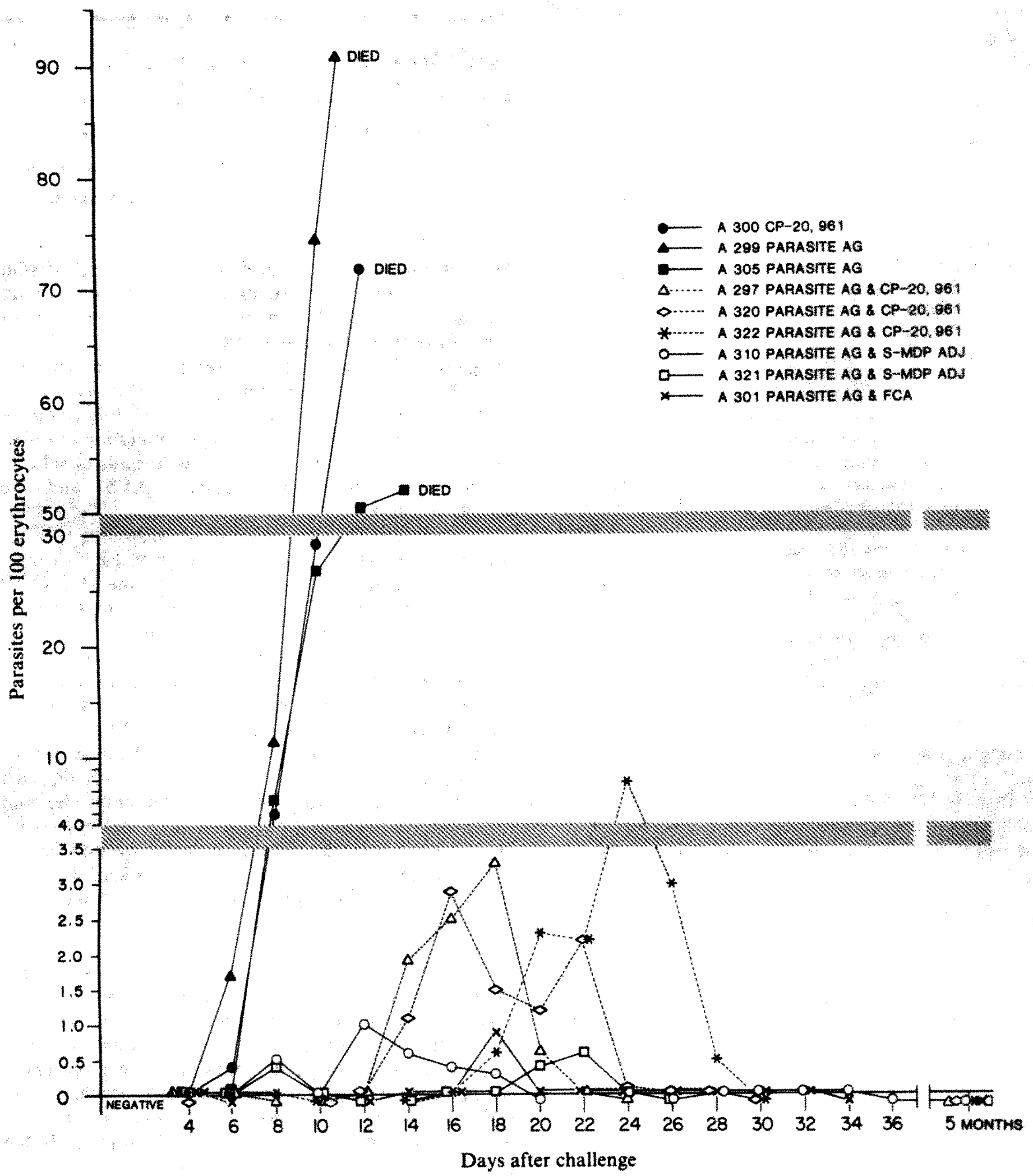


Fig. 1 Course of infection of *Plasmodium falciparum* (FUP strain) in control and vaccinated monkeys (*Aotus trivirgatus griseimembra*).

A297, A320 and A322 received vaccine composed of parasite antigen plus CP-20,961 adjuvant; A310 and A321 were vaccinated with parasite antigen mixed with stearoyl-MDP; A301 received parasite antigen with FCA. Two intramuscular injections of vaccine were given in alternate thighs at a 3-week interval. Immunization never produced a detectable infection. On day 51, all nine monkeys were challenged with an intravenous injection of 7.5×10^5 parasites (FUP strain of *P. falciparum*) derived from a current infection in an owl monkey. Thick and/or thin blood films were made daily to follow the course of infection (Fig. 1). All three control monkeys died within 2 weeks of challenge: A300 and A299 died on day 12 with parasitaemias of 72.0 and 91.0%, respectively, and A305 died on day 14 with 52.2% parasitaemia. In contrast, all six vaccinated monkeys survived challenge, exhibiting only low grade infections. The duration of patency was longest in A310, which experienced a maximum parasitaemia of 1.6% and a very mild recrudescence. The duration of parasitaemia was similar in the other five vaccinated animals. Peak parasitaemias of 4% occurred in A297 and A320 and 8% in A322, and lasted for 3 days only. In general, the level of parasitaemia remained below 0.5% in A301 and A321. Full recovery from *P. falciparum* infection began on day 24 for A297, on day 32 for A320, A321, A322 and A301, and on day 36 for A310. Although the number of monkeys used in this experiment was small, the difference between the course of infection in immunized and non-immunized monkeys is indeed very significant considering the high virulence of *P. falciparum* (FUP) for owl monkeys. The FCA-immunized monkeys displayed the usual granulomatous abscesses at the injection site, but only minimal side reactions were observed in the owl monkeys immunized with the synthetic adjuvants: stearoyl-MDP induced a temporary anorexia and CP-20,961 produced a slight, transient swelling at the site of injection.

Our results confirm previous findings of protective immunity conferred by merozoite vaccines containing stearoyl-MDP[11] and FCA[8,9], and further demonstrate that the novel synthetic adjuvant CP-20,961, when combined with *P. falciparum*

Table 1 Use of synthetic adjuvants in vaccination of Aotus monkeys against *Plasmodium falciparum* (Uganda-Palo Alto strain)

	Composition of vaccine		
Monkey no.	Adjuvant	Parasite antigen (1.4 mg protein per injection)	Challenge inoculum (7.5×10^5 infected erythrocytes)
A299	None	+	+
A305	None	+	+
A300	Pfizer, CP-20,961*	None	+
A297	Pfizer, CP-20,961	+	+
A320	Pfizer, CP-20,961	+	+
A322	Pfizer, CP-20,961	+	+
A310	S-MDP†	+	+
A321	S-MDP	+	+
A301	FCA‡	+	+

The vaccine consisted of adjuvant plus antigen mixed using a 3-way stopcock and two syringes, administered intramuscularly in alternate thighs. Vaccination was on days 0 and 21. Parasite antigen consisted of >50% mature schizonts with individual merozoites, freed by saponin lysis, 0.5 ml per dose[21]. Challenge inoculum was with parasitized blood obtained from a current infection (FUP strain of *P. falciparum*) in an *Aotus* monkey. Inoculum was given intravenously on day 51.

* Pfizer adjuvant containing 3 mg of CP-20,961 dispersed in Intralipid in 1.0 ml volume per dose[22].

† 0.25 mg of 5-*O*-stearoyl-*N*-acetylmuramyl-L-alanyl-D-isoglutamine plus liposomes in 1.0 ml volume per dose[21].

‡ 0.5 ml Freund's complete adjuvant (Difco) per dose.

merozoite antigen, fully protects owl monkeys against a normally lethal homologous challenge of intra-erythrocytic *P. falciparum* parasites. Although the duration of immunity conferred by malaria vaccines containing CP-20,961 or other adjuvants has not been fully assessed in the present model, antibodies against Rift Valley fever virus produced in *Macacca fascicularis* with a CP-20,961-adjuvanted vaccine persist at high levels for at least 6 months (J. A. Reynolds, personal communication).

This work was supported by AID contract ta-c-1227, UNDP/World Bank/WHO special program for research and training in tropical diseases, and Central Research Division, Pfizer Inc., USA.

Received 21 July; accepted 4 November 1980.

1. *WHO tech. Rep. Ser.* **579** (World Health Organization, Geneva, 1975).
2. Maugh, T. H. *Science* **196,** 413–416 (1977).
3. Young, M. D., Porter, J. A. Jr & Johnson, C. M. *Science* **153,** 1006–1007 (1966).
4. Geiman, Q. M. & Meager, M. J. *Nature* **215,** 437–439 (1967).
5. Geiman, Q. M. & Siddiqui, W. A. *Am. J. trop. Med. Hyg.* **18,** 351–354 (1969).
6. Trager, W. & Jensen, J. B. *Science* **193,** 673–675 (1976).
7. Siddiqui, W. A., Schnell, J. V. & Richmond-Crum, S. *Am. J. trop. Med. Hyg.* **23,** 1015–1018 (1974).
8. Siddiqui, W. A. *Science* **197,** 388–389 (1977).
9. Mitchell, G. H., Butcher, G. A., Richards, W. H. G. & Cohen, S. *Lancet* **ii,** 1335–1338 (1977).
10. Kotani, S. *et al. Biken's J.* **20,** 95–103 (1977).
11. Siddiqui, W. A. *et al. Science* **201,** 1237–1239 (1978).
12. Voller, A. & Richards, W. H. G. *Lancet* **ii,** 1172–1174 (1968).
13. Sadun, E. H., Wellde, B. T. & Hickman, R. L. *Milit. Med.* **134,** 1165–1175 (1969).
14. Wellde, B. T., Diggs, C. L. & Anderson, S. *Bull. Wld Hlth Org.* **57,** Suppl. 1, 153–157 (1979).
15. Siddiqui, W. A. *et al. Bull. Wld Hlth Org.* **57,** Suppl. 1, 199–203 (1979).
16. Reese, R. T., Trager, W., Jensen, J. B., Miller, D. R. & Tantravati, R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5665–5668 (1978).
17. Hoffman, W. W., Korst, J. J., Niblack, J. F. & Cronin, T. H. *Antimicrob. Ag. Chemother.* **3,** 498–502 (1973).
18. Glaz, E. T. & Talas, M. *Acta microbiol. acad. sci. hung.* **25,** 87–96 (1978).
19. Panusarn, C., Stanley, E. D., Dirda, V. Rubenis, M. & Jackson, G. G. *New Engl. J. Med.* **291,** 57–61 (1974).
20. Niblack, J. F. *Tex. Rep. Biol. Med.* **35,** 528–534 (1977).
21. Niblack, J. F. *et al. J. reticuloendothel. Soc.* **S26,** 655–666 (1979).
22. Chang, Y. H. & Pearson, C. M. *Arthritis Rheum.* **21,** 169–170 (1978).
23. Anderson, A. O. & Reynolds, J. A. *J. reticuloendothel. Soc.* **S26,** 667–680 (1979).
24. Siddiqui, W. A. & Schnell, J. V. *J. Parasit.* **59,** 516–519 (1973).
25. Siddiqui, W. A. & Richmond-Crum, S. M. *J. Parasit.* **63,** 583–584 (1977).
26. Siddiqui, W. A., Kramer, K. & Richmond-Crum, S. M. *J. Parasit.* **64,** 168–169 (1978).
27. Inoue, K. *Biochim. biophys. Acta* **339,** 390–402 (1974).
28. Ma, N. S. F., Jones, T. C., Miller, A. C., Morgan, L. M. & Adams, E. A. *Lab. Anim. Sci.* **26,** 1022–1036 (1976).
29. Moorehead, B. S., Nowell, P. C., Mellman, W. J., Battips, D. M. & Hungerford, D. A. *Expl Cell Res.* **20,** 613–616 (1960).
30. Taylor, W. & Siddiqui, W. A. *J. Parasit.* **65,** 267–271 (1979).

0028-0836/81/020066—03$01.00

Isolation of a cloned cell line expressing variant H–2K^k using fluorescence-activated cell sorting

B. Holtkamp, M. Cramer, H. Lemke* & K. Rajewsky

Institute for Genetics, University of Cologne, Weyertal 121,
D-5000 Köln 41, FRG

We describe here a method for the isolation of somatic cell variants that express a structurally altered gene product on their cell surface. The method makes it possible to select positively for cells of the variant phenotype which is defined by loss of a given antigenic determinant. As an example, we present the isolation of a cloned cell line expressing a structurally altered H–2K^k molecule, designated H–2K^{kS1}. Cells of this phenotype were found to be present in the wild-type population at a frequency of 10^{-5}–10^{-6} and were isolated in three rounds of selection in the fluorescence-activated cell sorter (FACS) and subsequent growth *in vitro*.

Our method for the isolation of structural H–2K^k variants requires a set of monoclonal antibodies, each of which reacts with a single antigenic determinant on the H–2K^k molecule[1]. The antigenic determinants recognized by these antibodies can be distinguished from each other by genetic analysis in which the antibodies are reacted with a set of cells expressing different H–2 haplotypes[2]. Imagine a pair of monoclonal anti-H–2K^k antibodies, each of which binds to a different H–2K^k determinant. If one antibody is labelled with a green and the other with a red fluorescent dye and the cells stained with a mixture of the two, we could in principle directly isolate, by cell sorting, variant cells that have lost reactivity with one of the antibodies but not the other. Such cells would be expected to continue expressing an H–2K^k-like molecule, but this molecule would lack at least one antigenic determinant. Because of technical limitations of the FACS I (Becton-Dickinson), we have not yet been able to apply this simple principle directly, but have used a similar, though less general, approach (Fig. 1).

We used a pair of monoclonal antibodies which mutually inhibit their binding to the H–2K^k target, presumably by steric hindrance. Such pairs of antibodies have been shown to exist[3]. One of the two antibodies is labelled with a fluorescent dye and incubated with cells in the presence of a large excess of the (unlabelled) second antibody. In these conditions we expect staining of variant cells only, in particular of variants that have lost reactivity with the second antibody. Such cells can then be positively selected for in the cell sorter.

The wild-type cells used in our study were descendants of the LDHB cell line which originated from a spontaneous lymphoma of a (C3H × DBA/2)F$_1$ mouse and which is described in ref. 3. LDHB cells possess 41 chromosomes and carry on one of their two chromosomes no. 17 the structural gene for the H–2K^k molecule. Because H–2K^k is only poorly expressed on the surface of LDHB cells[1], a 'high-producer' subline was established by staining LDHB cells with fluorescent, monoclonal anti-H–2K^k antibodies[2], isolating the brightest cells in the FACS I, growing them up in culture and repeating the process. After eight cycles of sorting and subsequent growth *in vitro*, the cells were cloned by limiting dilution (cloning efficiency 70%). The cells then grew in culture for 5 months with a doubling time of 18 h, during which time they were selected for high H–2K^k expression by cell sorting five times. The resulting cell line, HK13, exhibited bright fluorescence on staining with fluorescent anti-H–2K^k antibodies, 10 times above that of spleen cells from (C3H × DBA/2)F$_1$ mice, and from this cell line we isolated a somatic variant with structurally altered H–2K^k.

For this purpose we used fluorescein-labelled antibody H100-5/28 for cell staining and antibody H100-27/55 as inhibitor (compare with Fig. 1). The properties of these monoclonal

---

* Present address: Biochemical Institute, Christian-Albrecht-University, D-2300 Kiel, FRG.

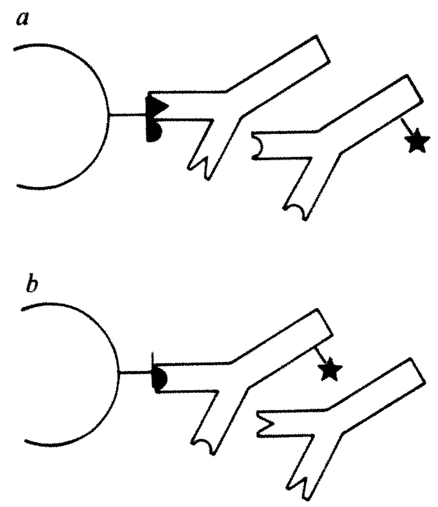


Fig. 1 Principle of variant selection. *a*, The labelling of cells carrying wild-type H-2K^k by a fluorescent (★) monoclonal antibody is inhibited by an unlabelled second antibody which binds to a neighbouring antigenic determinant. *b*, Cells expressing variant H-2K^k molecules that have lost the neighbouring determinant are successfully stained in the same experimental conditions. In reality, the unlabelled antibody is present in a 1,000-fold excess.

antibodies against H-2K^k are described in refs 1 and 2. The methods for attaching fluorescein groups to antibodies and for cell staining have been described previously[3]. Antibody H100-27/55 inhibited almost completely fluorescent staining of HK13 cells by antibody H100-5/28. Rare spontaneous variants whose staining could not be inhibited, and which were apparently present in a population of 2.4×10^7 HK13 cells (the total number of cells initially sorted; compare with Table 1), were isolated during three rounds of separation in the cell sorter and subsequent growth *in vitro*. The technical details of our cell-sorting procedure have been described previously[4]. Table 1 summarizes this particular series of cell sorting.

The proportion of variant (fluorescent) cells in the cell population increased dramatically after each sort (30–100-fold) and reached 78% after the third round of cell sorting and growth. These cells were again stained as before and variants were cloned directly by sorting single fluorescent cells into the wells of a 96-microwell tissue culture plate[5]. Fourteen clones were picked, recloned twice by limiting dilution (cloning efficiency ~70%) and subjected to fluorescence analysis. With one exception, all the clones behaved similarly in this analysis—almost 100% of the cells exhibited bright fluorescence when labelled with fluorescein-conjugated antibody H100-5/28, whereas the fluorescence on labelling with fluorescein-conjugated antibody H100-27/55 was close to background. (In fact the fluorescence was slightly above background. We attribute this to the slight cross-reaction of antibody H100-27/55 with the H-2D^k product[2] (also expressed in HK13 cells) because the staining could not be inhibited by unlabelled antibody H100-5/28.) The binding of antibody H100-5/28 to the variant cells was not significantly inhibited by even a 1,000-fold excess of antibody H100-27/55 in a radioinhibition assay[1] (data not shown).

To demonstrate that the variant cells indeed express structurally altered H-2K^k molecules, a clone exhibiting particularly strong fluorescence with antibody H100-5/28 was selected (HK13.S1) and grown up in tissue culture. 1×10^8 variant and wild type cells (HK13) were surface iodinated with ^{125}I by the lactoperoxidase technique[6] and lysed at 4 °C with 0.5% NP40 in the presence of 10^{-3} M phenylmethylsulphonyl fluoride. Nuclei were spun down and the lysate deaggregated by high-speed centrifugation. Details of the method are given in ref. 7. H-2K^k molecules were then precipitated from the lysates with monoclonal antibodies H100-27/55 and H100-5/28 and analysed on SDS-polyacrylamide gels (for experimental details see Fig. 2 legend). The results are shown in Fig. 2.

In the case of the HK13 cell lysate, the familiar 45,000- and 12,000-molecular weight bands of H-2K heavy and light chains appear on the gels on precipitation with either of the two monoclonal antibodies (tracks *a* and *b*). As staining of HK13 cells by either antibody can be completely inhibited by the other, we assume that the protein bands in tracks *a* and *b* represent the heavy and light chains of a single H-2K^k molecule that co-expresses the determinants recognized by the two antobodies. With the variant cell lysate, bands of similar size appear on the gel, but only on precipitation with antibody H100-5/28; anti-

Table 1 Protocol of H-2K^k variant enrichment

Sorting*	No. of cells sorted	% Cells deflected	% Fluorescent cells†
1st	2.4×10^7	1.2	0.025 (2×10^4)‡
2nd	2.9×10^7	2.0	0.65 (2.3×10^3)‡
3rd	2.1×10^7	0.6	77.5 (4×10^4)§

* Cells were incubated with 0.02 mg ml^{-1} fluorescein-labelled monoclonal antibody H100-5/28 and a 10^3-fold excess of unlabelled monoclonal antibody H100-27/55. Fluorescent cells were separated in the FACS.

† After sorting and subsequent growth *in vitro*. Values in parentheses indicate total number of cells analysed.

‡ Determined by microscopic analysis.

§ Determined by cell-sorter analysis.

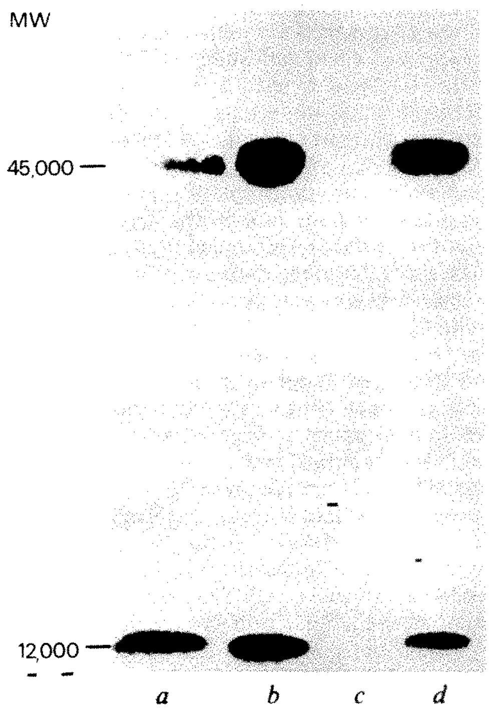


Fig. 2 Electrophoretic analysis of precipitations of wild-type and variant H-2K^k molecules with monoclonal antibodies H100-27/55 and H100-5/28. Detergent lysates (equivalent to 10^7 surface-iodinated cells) were incubated with 30 μg antibody (DEAE-cellulose purified) at 4 °C for 6 h and the immunoglobulin (IgG2a in both cases) was bound to 25 μl protein A-Sepharose (Pharmacia). After appropriate washings; sample buffer 4% SDS, 6 M urea, 10% 2-mercaptoethanol and bromophenol blue (tracking dye) was used to dissociate the antigen-antibody-protein A-Sepharose complexes. The figure shows the autoradiograph (8 days exposure) of a 10–15% gradient polyacrylamide slab gel run in 0.1% SDS in the discontinuous system of Laemmli[10]. One aliquot of HK13 (wild-type) lysate was precipitated with antibody H100-27/55 (track *a*) and another aliquot was precipitated with antibody H100-5/28 (track *b*). Similarly, HK13.S1 lysate (variant type) was precipitated with antibody H100-27/55 (track *c*) and with antibody H100-5/28 (track *d*). As identical amounts of antibodies H100-27/55 and H100-5/28 were used in our precipitating procedure, we consider the fainter bands in track *a* to be due to a lower precipitation capacity of antibody H100-27/55. MW, molecular weight.

body H100-27/55 does not bring down any detectable radioactivity (tracks *c* and *d*). Taken together, the data show that HK13.S1 cell membranes contain variant H-2K^k molecules, whose polypeptide chains are of similar size to those of the wild-type molecules but which lack one of the two antigenic determinants recognized by our two monoclonal antibodies and are co-expressed in wild-type H-2K^k. The different fluorescent-staining patterns of wild-type and variant cells are thus due to an alteration of H-2K^k itself and not, for example, to a change in its orientation in the cell membrane.

The selection method described here makes it possible to isolate efficiently and at will a somatic-cell variant that stably expresses structurally altered H-2K^k molecules on the surface. From the data in Table 1, we conclude that spontaneous variants of the selected phenotype were present in the population of HK13 cells at a frequency of 10^{-5}–10^{-6}. Taking into account the generation time of the cells and the number of generations after cloning, we estimated roughly a frequency of variation of 10^{-6}–10^{-7} per cell per generation. This low frequency and the stability of the variant suggest that the phenotypic change in HK13.S1 cells is due to mutation. Preliminary experiments indicate that both antibodies used in the variant selection also react with unglycosylated H-2K^k molecules. The mutation carried by HK13.S1 cells may therefore have occurred in the structural gene for H-2K^k. Structural analysis of the variant protein or of the corresponding genetic material is required to establish this point. We provisionally designate the mutation as H-2K^{ks1} (the prefix 's' standing for 'somatically derived'), assuming that the mutation is expressed exclusively in the H-2K^{kS1} molecule. We have meanwhile isolated a second H-2K^k structural variant using a different pair of monoclonal antibodies (unpublished data) and thus feel that our method of selection may have wide applicability.

The advantages of our approach are that many cells can be screened and that we exclude variants that have lost the expression of the entire surface component in question—an event which is often more frequent than structural variation[8]. The latter feature also applies to the classical identification of structural H-2 mutations carried in the germ line by tissue transplantation between inbred animals (see ref. 9) but which is severely limited in quantitative terms.

The present approach allows us to study not only types and frequencies of mutations occurring in mammalian genes, such as genes in the H-2 complex and V and C genes encoding immunoglobulin heavy chains, but also the functional properties of the variant-gene products. With H-2 this might elucidate the role of H-2-encoded molecules in the immune system. It is encouraging that we have found that the variant H-2K^k molecules expressed in HK13.S1 cells function very efficiently as targets for alloreactive killer T cells but are deficient as restricting elements in trinitrophenyl-specific, H-2-restricted cytotoxic reactions (J. Neuerburg and H.-W. Vohr, unpublished data).

This work was supported by the Deutsche Forschungsgemeinschaft through SFB 74 and the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen. We thank W. Santüns for help in the cell sorter experiments and Ms C. Schenkel for technical assistance, Dr K. Reske for technical advice and Dr A. Radbruch for analysing the LDHB cell karyotype.

Received 2 September; accepted 7 November 1980.

1. Lemke, H., Hämmerling, G. J. & Hämmerling, U. *Immun. Rev.* **47,** 175–206 (1979).
2. Klein, J., Huang, H. S., Lemke, H., Hämmerling, G. J. & Hämmerling, U. *Immunogenetics* **8,** 419–432 (1979).
3. Holtkamp, B., Lindahl, K. F., Segall, M. & Rajewsky, K. *Immunogenetics* **9,** 405–421 (1979).
4. Liesegang, B., Radbruch, A. & Rajewsky, K. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3901–3905 (1978).
5. Parks, D. R., Bryan, V. M., Oi, V. T. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1962–1966 (1979).
6. Haustein, D. J. *Immun. Meth.* **27,** 25–38 (1975).
7. Jones, P. P. *J. exp. Med.* **146,** 1261–1279 (1977).
8. Rajan, T. V. *Immunogenetics* **10,** 423–431 (1980).
9. Klein, J. *Adv. Immun.* **26,** 55–146 (1978).
10. Laemmli, U. K. *Nature* **227,** 680–685 (1970).

Expression of cell-surface HLA-DR, HLA-ABC and glycophorin during erythroid differentiation

Jean Robinson*, Colin Sieff†, Domenico Delia*, Paul A. W. Edwards‡ & Melvyn Greaves*

* Membrane Immunology Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK
† Department of Haematology, The Hospital for Sick Children, Great Ormond Street, London WC1, UK
‡ Ludwig Institute for Cancer Research, Sutton, Surrey SM2 5PX, UK

The unexpected discovery that Ia-like (HLA-DR) antigens[1-3] in humans were present on blast cells from acute myeloblastic leukaemia[4,5] led to the finding that normal granulocytic progenitors, in contrast to their mature descendents, also expressed HLA-DR antigens. Thus, anti-Ia sera stain a proportion of myeloblasts in normal bone marrow[6,7], inhibit myeloid progenitor (CFU-GM) colony formation in the presence of complement[8] and can be used to label and separate CFU-GM on a fluorescence-activated cell sorter (FACS)[9]. Winchester *et al.* subsequently reported that erythroid progenitors (BFU-E and CFU-E) were also inhibited or killed by anti-Ia (p28,37) and complement[10]. These observations raised the possibility that HLA-DR (or presumptive I-region equivalent) products might have a regulatory role in early haematopoiesis. We have now analysed HLA-DR and HLA-ABC antigen expression on normal erythroid progenitors using monoclonal antibodies to non-polymorphic determinants and fluorescence-activated cell sorting. In parallel experiments, we tested a monoclonal antibody to glycophorin, a well defined erythroid-specific cell-surface membrane glycoprotein[11]. We report that HLA-DR, HLA-ABC and glycophorin are all expressed at various stages during erythroid differentiation.

Table 1 lists the results of assays for erythroid burst-forming units (BFU-E) and colony-forming units (CFU-E) in a series of experiments in which monoclonal anti-HLA-DR was used to stain blood or marrow cells followed by analysis and sorting on the FACS (Fig. 1).

The majority of recoverable BFU-E consistently seem to localize in the HLA-DR-positive population, whereas CFU-E are found mostly, but not exclusively, in the fraction which has either low density or no HLA-DR. When light scattering was used as an additional parameter for sorting, both the BFU-E and CFU-E seemed to be restricted to the population within the larger light-scattering (or presumed bigger) cell size range.

The total recovery of BFU-E in the separated fractions was, however, only approximately one-third (Table 1 legend). Much higher recovery of CFU-E in these experiments and of BFU-E in experiments with anti-HLA-ABC and anti-glycophorin indicate that this is probably not a technical artefact; neither is it attributable to an inhibiting effect of the monoclonal anti-HLA-DR because no diminution of BFU-E activity is seen in the stained but unsorted cells (see control 2 in Table 1). This result may therefore indicate that sorting on the basis of HLA-DR expression separates cells whose interaction may be required for optimal BFU-E activity.

When monoclonal anti-HLA-ABC (PA 2.6) was used to stain bone marrow cells, all BFU-E and CFU-E were sorted into the HLA-ABC$^+$ (positive) population (Table 1, experiments 6, 7). Bone marrow cells were sorted on the basis of intensity of fluorescence staining with monoclonal anti-glycophorin (Fig. 1). More than 99% of both BFU-E and CFU-E were found in the fraction giving minimal or negative staining (Table 1, experiments 8, 9).

Less than 5.0% of the morphologically recognizable erythroid cells were separated into the HLA-DR-positive fractions. By contrast, almost all the cells in the glycophorin-positive popu-

0028-0836/81/020068—04$01.00

Table 1 Separation of BFU-E and CFU-E with monoclonal antibody and the FACS

				Colony assays					
				CFU-E			BFU-E		
Expt	Antibody	Fraction	% Cells	Per 10^5 cells	Recovery per 10^5 unfractionated cells (%)	% Of total recovered†	Per 10^5 cells	Recovery per 10^5 unfractionated cells (%)	% Of total recovered
1	DA2 (HLA-DR)	1. Control					23		
		2. Control					18		
		Sorted:							
		3. +Large	30				14	24.5	95.3
		4. +Small					149		
		5. Negative	60				2	1.2	4.7
								25.7(125)*	
2	DA2 (HLA-DR)	1. Control					15.8		
		2. Control					8.3		
		Sorted:							
		3. +Large	10				19.4	1.94	100
		4. +Small	6.5				0	0	0
		5. Negative	69				0	0	0
								1.94(16)	
3	DA2 (HLA-DR)	1. Control		38.8			12.5		
		2. Control		46.8			14.3		
		Sorted:							
		3. +Large	3	288	8.6	24	133	3.99	99.5
		4. +Small	1	33	0.3	0.9	2	0.02	0.5
		5. Negative	96	27	25.9	74.4	0	0	0
					34.8 (81)			4.01(30)	
4	DA2 (HLA-DR)	1. Control		81.5			21.8		
		2. Control		120.5			19		
		Sorted:							
		3. +Large	9	94	8.5	14.9	56	5.04	89.4
		4. +Small	10	10.5	1.0	1.8	1	0.14	1.8
		5. Negative	50	95	47.5	83.0	1	0.50	8.8
					57.0 (56)			5.64(28)	
5	DA2 (HLA-DR)	1. Control		445			163		
		2. Control		489			217		
		Sorted:							
		3. +Bright	9	94	8.5	3.3	109	9.8	16.0
		4. +Dim	10	676	67.6	26.0	400.5	40.0	65.3
		5. Negative	74	248	183.5	70.7	15.5	11.5	18.7
					259.6 (56)			61.3(32)	
6	PA 2.6 (HLA-ABC)	1. Control		210			84		
		2. Control		302			114		
		Sorted:							
		3. +Bright/dim	42	258	108	92	121	50.8	93
		4. Negative	32	27.5	8.9	8	12	3.8	7
					117 (45)			54.7(55)	
7	PA 2.6 (HLA-ABC)	1. Control		312			42.5		
		2. Control		313			63.5		
		Sorted:							
		3. +Bright/dim	61	410	250.0	97.3	199	121.4	99.6
		4. Negative	22	32	7.05	2.7	2	0.4	0.3
					257.05 (82)			121.8(232)	
8	28/10.1 (Glycophorin)	1. Control		112.5			94		
		2. Control		123.5			106.5		
		Sorted:							
		3. +Bright	3.4	44.4	1.4	1.0	35.6	1.05	1.1
		4. +Dim		40.6			26.3		
		5. Negative	96.6	150	144.9	99	100.5	97.1	98.9
					146.3 (123)			98.15(98)	
9	28/10.1 (Glycophorin)	1. Control		371			137		
		2. Control		321			112		
		Sorted:							
		3. +Bright	5	31.4	1.6	0.5	2.8	0.14	0.14
		4. +Dim	8	98.5	7.9	2.5	15.5	1.24	1.26
		5. Negative	78	399.5	311.6	97.0	123.5	96.3	98.6
					321 (93)			97.7(79)	

The unfractionated controls and the positive and negative cell fractions were washed three times in cold alpha medium (Flow) and 2% fetal calf serum (FCS). The cells were then cultured in a mixture containing 30% FCS (Gibco Bio-Cult), 1% bovine serum albumin (Sigma), 10^{-4} M mercaptoethanol, penicillin, streptomycin, 0.9% methylcellulose (Dow) and 2.5 units ml^{-1} erythropoietin (Connaught Step III)[24]. The final concentration of nucleated cells was 1 or 2×10^5 ml^{-1}, although occasionally $<1\times10^5$ cells were obtained for some of the fractions. Volumes of 1 ml of the mixture were plated in 35-mm tissue culture plates (Nunc) in duplicate, and incubated at 37 °C in a high humidity 5% CO_2/95% air tissue-culture incubator. Recognizable CFU-E, identified by the pink-red colour of the cells and comprising one or two clusters, were counted directly on day 7 using an inverted microscope at ×40 magnification. BFU-E, large, red colonies comprising three or more sub-colonies, were counted on day 14. In each experiment, controls 1 and 2 are unlabelled and labelled fractionated cells, respectively. Recovery per 10^5 unfractionated cells is determined by multiplying the number of colonies per 10^5 cells (column 2) by the proportion of cells in that fraction.

* Total colonies recovered, expressed (in parentheses) as a percentage of the mean of the two unfractionated controls. Note, however, that selection of cells on the FACS involves some loss of cells.

† Number of colonies recovered in different fractions expressed as a percentage of total colonies recovered.

lation were identifiable as erythroid precursors at different stages of maturation; almost no recognizable erythroid cells were found in the glycophorin-negative fraction.

It has been suggested that BFU-E activity is enhanced by T cells[12]. As most T cells have no detectable glycophorin and are largely though not entirely HLA-DR negative, potential T-cell–BFU-E interactions could be compromised by the cell-separation experiments described above. To investigate this possibility we labelled bone marrow T cells with a mixture of two pan-T monoclonal antibodies OKT1 (ref. 13) and OKT11 (ref. 14), separated stained and unstained cells on the FACS and assessed the effect of recombination of these fractions. The results (Table 2) indicate that effectively all BFU-E and CFU-E are unreactive with these particular monoclonal anti-T reagents and that bone marrow T cells have, if anything, an inhibitory effect on BFU-E at the cell concentrations used. Further experiments are required to investigate this possibility more thoroughly. Note, however, that the reported promoting effect of blood T cells on BFU-E may be restricted to circulating erythroid progenitors[12] and that this response has not been confirmed in another laboratory[15].

Fluorescence-activated cell sorting with monoclonal antibodies provides an incisive approach to the analysis of cell-surface antigen expression on numerically infrequent progenitor cells. Our data using this approach suggest that the earliest detectable, committed erythroid progenitors, BFU-E, express HLA-DR as well as HLA-ABC antigens and that the former are progressively lost during differentiation through CFU-E to morphologically recognizable erythroblasts. HLA-ABC antigens are lost more slowly, for the majority of CFU-E are positive, as are at least a proportion of morphologically recognizable precursors. Although mature red blood cells have no detectable HLA-ABC, previous studies have indicated that normoblasts and reticulocytes may express at least a low density of these antigens[16–18].

In contrast, the antigenic determinant of glycophorin detected by the monoclonal antibody we have used is not expressed until the post-CFU-E, HLA-DR-negative stage of erythroid maturation. More than 90% of all morphologically recognizable erythroblasts were sorted into the glycophorin$^+$, HLA-DR$^-$ fraction on the FACS. Gahmberg *et al.* similarly reported that basophilic normoblasts had a low density of cell-surface glycophorin but could not see binding of rabbit antibodies to glycophorin on less mature pronormoblasts[19].

We incidentally noted that CFU-GM (which grow in colony assays for erythroid cells) separate predominantly in the HLA-DR$^+$, HLA-ABC$^+$, glycophorin$^-$ fractions. This confirms previous cell-sorting experiments in which CFU-GM were present in the fraction stained with non-monoclonal anti-p28,33 (Ia-like or HLA-DR antigen)[9].

Table 2 Effect of bone marrow T cells on BFU-E and CFU-E

		CFU-E per 10^5 cells			BFU-E per 10^5 cells		
Expt	Fraction	Obs.	Exp.*	Obs./Exp.	Obs.	Exp.	Obs./Exp.
(1)	1. Control	182			72		
	Sorted:						
	2. Negative	406			266		
	3. Bright$^+$	4.5			2.5		
	4. Dim$^+$	36.5			21.7		
	Sorted and recombined:						
	5. Negative + bright$^+$	148.5	132	1.1	47.5	86.3	0.55
	6. Negative + dim$^+$	154	148	1.0	99.5	96.3	1.0
(2)	1. Control	415			144		
	Sorted:						
	2. Negative	548			302.5		
	3. Bright$^+$	66			12.5		
	4. Dim$^+$	104			29.5		
	Sorted and recombined:						
	5. Negative + bright$^+$	330	299	1.1	123.5	156	0.8
	6. Negative + dim$^+$	331	326	1.0	115	166	0.7

Bone marrow cells stained with optimal concentration of OKT1 plus OKT11, sorted on the FACS and plated *in vitro* for CFU-E and BFU-E assays (see Table 1 legend). Control = stained but unseparated. Expt 1.5: negative (non-T)/bright (T) ratio 1:1. Expt 1.6: negative (non-T)/dim (T) ratio 0.6:1 (insufficient cells available for 1:1 ratio). Expt 2.5: negative (non-T)/bright (T) ratio 1:0.8 (insufficient cells available for 1:1 ratio). Expt 2.6: negative (non-T)/dim (T) ratio 1:1.

* Calculated for the number of cells added to mixtures and the ratio of CFU-E to BFU-E per 10^5 cells in each separated population before mixing.

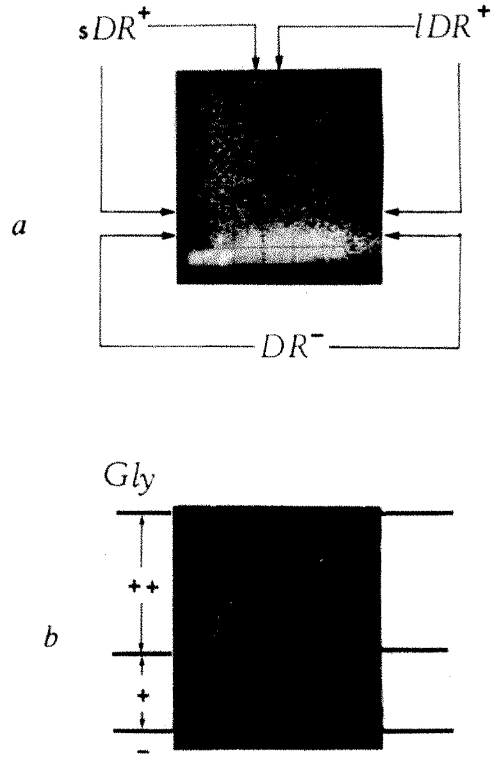


Fig. 1 Separation of erythroid precursors using monoclonal anti-HLA-DR and the FACS. Normal bone marrow was obtained by aspiration from the posterior iliac crest of adult human volunteers, or during femoral osteotomy of haematologically normal children. Approximately 5–10 ml of marrow from two to four aspirations were collected into 10 ml Hank's buffered salt solution containing 100 U of preservative-free heparin. Marrow suspensions were centrifuged (400*g* for 40 min at 20 °C), in Ficoll–Isopaque (1.077 g cm^{-3}), the interface mononuclear cells collected, washed twice and suspended in Eagle's medium plus 20% FCS. Murine monoclonal antibodies were produced using cell hybridization methods as described previously. Monoclonal anti-HLA-DR (DA2) identifies a non-polymorphic determinant associated with an Ia-like polypeptide (p28,34) complex on B lymphoblasts[22]. Monoclonal antibody to HLA-ABC (PA 2.6) similarly recognizes a non-polymorphic determinant shared by products of the A, B, C but not the D locus[22]. Monoclonal antibody LICR.LON/R10 was raised against erythrocytes and specifically precipitates glycophorin A[23]. 20–100 × 10^6 bone marrow cells in 50–100 μl were incubated for 30 min at 4 °C with previously determined optimal concentrations of the various monoclonal antibodies. After two washes, the cells were stained with an affinity-purified $F(ab')_2$ preparation of goat antibodies which had been cross-absorbed with insolubilized human immunoglobulin and labelled with fluorescein isothiocyanate. The stained cell suspension was analysed by FACS (modified FACS-I, Becton Dickinson)[9]. Cells were separated in sterile conditions using relative fluorescence intensity (vertical axis, *a*, *b*) and in some experiments relative light-scattering levels (horizontal axis) to separate cells according to size (*b*) (sDR$^+$ = small HLA-DR-positive cells; lDR$^+$ = large HLA-DR-positive cells; DR$^-$ = total HLA-DR-negative cells; Gly = glycophorin; ++, brightly stained with anti-glycophorin; +, dimly stained; −, unstained).

Winchester *et al.*[10] previously reported that both BFU-E and CFU-E could be inhibited by treatment with a rabbit anti-p28,37 (Ia-like) serum and rabbit complement before colony assay. This observation suggests that erythroid progenitors express Ia-like (HLA-DR) cell-surface antigens and/or that other cells essential for BFU-E and CFU-E activity are inhibited by the antisera used. The experiments reported above demonstrate unequivocally that some BFU-E express a monomorphic HLA-DR antigenic determinant but that CFU-E were predominantly HLA-DR negative. Anti-p28,33 (Ia-like) antisera may have exceedingly high cytotoxic titres which are in excess of binding titres as assessed by immunofluorescence[11]. It is therefore possible that CFU-E have a very low density of HLA-DR antigens which was below our 'selected' level of positivity on the FACS but sufficient to render the cells susceptible to antibody-dependent lysis (or functional inhibition) by complement.

Pluripotent haematopoietic stem cells seem to be Ia-antigen (HLA-DR equivalent) negative in mice[20], and possibly also in man[21]. The transient expression of HLA-DR on early erythroid and granulocytic lineage progenitors suggests that cell-surface molecules carrying these determinants might be important in cell interactions regulating committed progenitor cells. Note, finally, that as there is no evidence for biosynthesis of HLA-DR by these progenitor cells, we cannot exclude the possibility that these cell-surface-associated molecules are somehow 'acquired' from other cells in the bone marrow.

This work was supported in part by the Imperial Cancer Research Fund and Child Health Research Appeal Trust. We thank Dr W. Bodmer for supplying monoclonal antibodies to HLA-DR and HLA-ABC and Drs G. Goldstein and P. Kung (Ortho Pharmaceuticals) for monoclonal anti-T reagents (OKT1, OKT11).

Received 4 July; accepted 21 October 1980.

1. Moller, G. *Transplantn Rev.* **30,** (1976).
2. Allison, J. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3952 (1978).
3. Barnstable, C. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **41,** 443 (1977).
4. Schlossman, S. F., Chess, L., Humphreys, R. E. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73,** 1288 (1976).
5. Greaves, M. F. & Janossy, G. *Biochim. biophys. Acta* **516,** 193 (1978).
6. Winchester, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **74,** 4012 (1977).
7. Janossy, G. *et al. Br. J. Haemat.* **37,** 391 (1977).
8. Koeffler, H. P., Niskanen, E., Cline, M., Billing, R. & Golde, D. *Blood* **54,** 1188 (1979).
9. Janossy, G., Francis, G., Capellaro, D., Goldstone, A. & Greaves, M. F. *Nature* **276,** 176 (1978).
10. Winchester, R. J. *et al. J. exp. Med.* **148,** 613 (1978).
11. Marchesi, V. T., Furthmayr, H. & Tomita, M. *A. Rev. Biochem.* **45,** 667 (1976).
12. Nathan, D. *et al. J. exp. Med.* **147,** 324 (1978).
13. Kung, P. C., Goldstein, G., Reinherz, E. L. & Schlossman, S. F. *Science* **206,** 347 (1979).
14. Greaves, M. F. *et al.* in *Modern Trends in Human Leukaemia* IV (eds Neth, R., Gallo, R. C., Hofschneider, P.-H. & Mannweiler, K.) (Springer, in the press).
15. Nomdedeu, B. *et al. Expl Hemat.* **8,** 854 (1980).
16. Brown, G., Biberfeld, P., Christensson, B. & Mason, D. Y. *Eur. J. Immun.* **9,** 272 (1979).
17. Kourilsky, F. M. *et al. J. Immun.* **106,** 454 (1971).
18. Harris, R. & Zervas, J. D. *Nature* **221,** 1062 (1969).
19. Gahmberg, C. G., Jokinen, M. & Andersson, L. C. *Blood* **52,** 379 (1978).
20. Basch, R. S., Janossy, G. & Greaves, M. F. *Nature* **270,** 520 (1977).
21. Moore, M. A. S. *et al. Blood* **55,** 682 (1980).
22. Brodsky, F. M., Parham, P., Barnstable, C. J., Crumpton, M. J. & Bodmer, W. F. *Immun. Rev.* **47,** 3 (1979).
23. Edwards, P. A. W. *Biochem. Soc. Trans.* **8,** 334 (1980).
24. Iscove, N. N. & Sieber, F. *Expl Haemat.* **3,** 32 (1975).

Tetrapeptides of the eosinophil chemotactic factor of anaphylaxis (ECF-A) enhance eosinophil Fc receptor

M. Capron, A. Capron, E. J. Goetzl & K. F. Austen

Centre d'Immunologie et de Biologie Parasitaire, Institut Pasteur,
59019 Lille Cedex, France
Howard Hughes Medical Institute Laboratory at Harvard Medical
School, Boston, Massachusetts 02115
Robert B. Brigham Hospital, Harvard Medical School, Seeley G.
Mudd Bldg, 250 Longwood Ave, Boston, Massachusetts 02115

Certain mast cell products, including the tetrapeptides Ala-Gly-Ser-Glu (Ala^4) and Val-Gly-Ser-Glu (Val^4) of the eosinophil chemotactic factor of anaphylaxis (ECF-A)[1] and histamine[2], are preferentially chemotactic for eosinophils *in vitro*. Amino acid substitutions or modifications of the ECF-A tetrapeptides may substantially alter the chemotactic activity of human eosinophils[3]. The cellular mechanism whereby ECF-A tetrapeptides enhance IgG-dependent killing of schistosomula by rat eosinophils[4] is, however, unclear. We now report that the parent tetrapeptides and the substituted analogue Val-Pro-Ser-Glu (Pro^3), but not histamine, increase the number of rat and human eosinophils that form rosettes with erythrocytes bearing IgG antibodies. Moreover, we find a correlation between the effect of these substances on the expression of rat eosinophil IgG Fc receptors and their capacity to increase the IgG-dependent cytotoxicity of rat eosinophils for *Schistosoma mansoni* schistosomula.

Rat eosinophil-rich populations were composed of non-adherent peritoneal cells from normal Fischer rats[4]. Rat peritoneal neutrophils were obtained 7 h after the intraperitoneal injection of 10% proteose-peptone in saline. Human eosinophils and neutrophils were obtained from venous blood of normal subjects or patients with hypereosinophilia not due to parasitic infections and prepared by centrifugation on metrizamide discontinuous gradients[5]. The different leukocyte populations were adjusted to 4×10^6 ml^{-1} in minimal essential medium. Various concentrations of the agonists or medium alone were added to the leukocyte populations and the mixtures were incubated under a CO_2 atmosphere at 37 °C with constant agitation. Immunoglobulin Fc receptors were estimated by the EA-rosette technique using sheep erythrocytes (E) sensitized with the IgG fraction of antisera to E prepared in rats (EA rat) or in rabbits (EA rabbit)[6] (Fig. 1).

The percentage of control rat eosinophils that formed EA rosettes was significantly higher with EA rabbit than with EA rat (Fig. 1), perhaps implying a higher affinity of the rat eosinophil Fc receptor for rabbit immunoglobulins. (Similar differences with immunoglobulins from different animal species have already been noticed with guinea pig eosinophils[7].) When rat eosinophils were pretreated with increasing concentrations of Val^4, Ala^4 and Pro^3 tetrapeptide, there was a dose-dependent increase in the percentage of eosinophils that formed EA rat rosettes. Neither Val-Gly-Ser-Glu-methyl ester nor the lower concentrations of histamine significantly altered the expression of rat IgG Fc receptors on eosinophils, but 10^{-4} M histamine depressed receptor expression (Fig. 1*a*). The lowest concentration of three of the peptides and all concentrations of the Val-Gly-Ser-Glu-methyl ester and histamine examined suppressed the expression of receptors for EA rabbit relative to

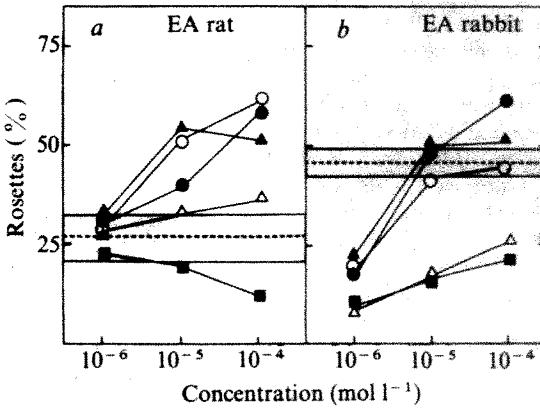


Fig. 1 Role of increasing concentrations of the ECF-A tetrapeptides Ala-Gly-Ser-Glu (●) (Serva) and Val-Gly-Ser-Glu (○), the analogues Val-Pro-Ser-Glu (▲) and Val-Gly-Ser-Glu-methyl ester (△) and histamine (■) (Sigma) on EA rat and EA rabbit rosette formation by rat eosinophils after 60 min incubation at 37 °C. A 60-min incubation was used for all subsequent experiments, as a preliminary study on the influence of incubation time has shown that the maximum effect occurs after 60 min incubation. Abscissa: final concentration of the products. The dotted line represents the mean percentage rosettes of rat eosinophils incubated for 60 min with minimal essential medium (±s.e.). The γ-globulin (IgG) fraction of rat or rabbit antisera to sheep erythrocytes (E) was prepared as described in ref. 6. Rat IgG-coated E (EA rat) or rabbit IgG-coated E (EA rabbit) were prepared by mixing equal volumes of rat or rabbit IgG anti-E with a 2% solution of fresh E as described[6]. Mixtures of 0.1 ml of the leukocyte suspensions and 0.1 ml of E, EA rat or EA rabbit at a concentration of 3×10^8 ml^{-1} were centrifuged at 100*g* for 10 min at 4 °C and incubated for 30 min at 0 °C. The pellet was gently resuspended and cytocentrifuge preparations were made. These were fixed in methanol and treated with Giemsa stain. All leukocytes with three or more adherent EA were considered to be rosettes. The percentage of 300 eosinophils that formed rosettes was counted for each slide. All results are expressed as the mean of the values for triplicate studies. Val-Gly-Ser-Glu, Val-Pro-Ser-Glu and Val-Gly-Ser-Glu-methyl ester were prepared according to ref. 3. Highly purified peptides were prepared by standard solid phase method and subsequently purified by sequential application of gel filtration on Sephadex G-25 in 0.05 M acetic acid and then reverse phase HPLC on a C18 column using sodium phosphate 0.02 M, *p*H 4.2. Thus, no active fragment of endotoxin could be present. Each point represents the mean of at least five experiments. Data were compared by one-way analysis of variance ($P < 0.05$) and Student's *t*-test. The increase in EA rat rosettes was significant for Ala-Gly-Ser-Glu (10^{-4} mol l^{-1}) and Val-Gly-Ser-Glu and Val-Pro-Ser-Glu (10^{-4} and 10^{-5} mol l^{-1}) ($P < 0.05$, *a*). Ala-Gly-Ser-Glu significantly increased the % of EA rabbit rosettes at the 10^{-4} mol l^{-1} concentration ($P < 0.05$, *b*).

0028-0836/81/020071—03$01.00

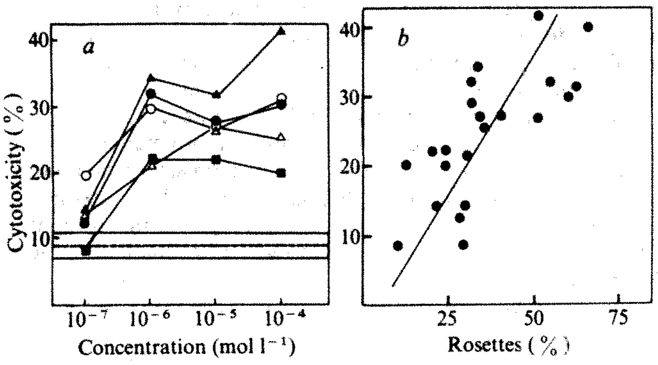


Fig. 2 *a*, Effect of increasing concentrations of Ala-Gly-Ser-Glu (●), Val-Gly-Ser-Glu (○), Val-Pro-Ser-Glu (▲), Val-Gly-Ser-Glu-methyl ester (△) and histamine (■) on rat IgG-dependent eosinophil cytotoxicity for schistosomula *in vitro*. Abscissa: final concentration of the products in the cytotoxicity assay. Ordinate: percentage dead schistosomula after 48 h contact with effector cells. *S. mansoni* schistosomula (prepared by penetration of cercariae through mice abdominal skin) were incubated overnight with heat-inactivated sera from *S. mansoni* infected rats[4]. Eosinophil-rich cell populations that were initially composed of 49% eosinophils and 10% mast cells were depleted of mast cells by centrifugation on 22.5% metrizamide[14] to yield a population of 53% eosinophils and 0.5% mast cells. The purified eosinophils were added to antibody-coated schistosomula at an effector to target ratio of 6,000 : 1, in a total volume of 200 μl. Aliquots (20 μl) of the various mediators were then added to each well of the cytotoxicity assay, giving final concentrations of 10^{-7}–10^{-4} mol l^{-1} (ref. 4). All values represent the mean of 3–13 experiments. Data were compared by one way analysis of variance ($P < 0.01$) and Student's *t*-test. The percentage of cytotoxicity was measured by direct visual examination of dead schistosomula in triplicate experiments[4]. The dotted line represents the mean percentage of cytotoxicity ± s.e. of antibody-coated schistosomula incubated with the mast cell-depleted eosinophils in medium alone. Ala-Gly-Ser-Glu, Val-Gly-Ser-Glu and Val-Pro-Ser-Glu significantly increased the levels of cytotoxicity for the three highest concentrations ($P < 0.05$). The increase with Val-Gly-Ser-Glu-methyl ester is only significant for the 10^{-5} mol l^{-1} concentration, and no significant effect was obtained with histamine. *b*, Correlation between the percentage of rat eosinophils that formed EA rat rosettes (abscissa) and the IgG-dependent cytotoxicity of rat eosinophils for schistosomula (ordinate), in the presence of the different mediators that had been tested (see Fig. 1*a* and *a* above). The effect of different concentrations of the mediators on the expression of rat IgG Fc receptors is significantly correlated with the increase in IgG-dependent rat eosinophil cytotoxicity for *S. mansoni* schistosomula ($r = 0.72$, $P < 0.001$; $n = 23$). A statistical analysis showed best fitting with a first-order regression line.

the levels detected on control eosinophils. Only 10^{-4} M Ala-Gly-Ser-Glu significantly increased the apparent number of EA rabbit rosettes (Fig. 1*b*).

The experimental model previously used to show that the ECF-A tetrapeptides could stimulate rat IgG-dependent eosinophil cytotoxicity for *S. mansoni* schistosomula *in vitro*[4] was applied in parallel with the EA-rosette assay. A population of normal rat peritoneal cells enriched in eosinophils and depleted of mast cells was incubated with various concentrations of the mast cell mediators and rat IgG antibody-coated schistosomula (Fig. 2*a*). A highly significant increase in the levels of cytotoxicity was obtained with the ECF-A tetrapeptides Ala-Gly-Ser-Glu and Val-Gly-Ser-Glu and with the analogue Val-Pro-Ser-Glu at the three highest concentrations. A smaller effect was detected with Val-Gly-Ser-Glu–methyl ester, and histamine showed no significant effect. Comparison of the EA rat rosette and the IgG-dependent cytotoxicity assay results as a function of the dose of the tetrapeptides or histamine revealed a highly significant correlation ($r = 0.72$, $P < 0.001$; Fig. 2*b*).

To determine whether the Fc receptor enhancement was restricted to rat eosinophils, similar experiments were performed with human eosinophils. When blood eosinophils from patients with parasite-unrelated hypereosinophilic syndromes were preincubated with various mediators, an increase in the percentage of rosettes was obtained both with EA rat and with EA rabbit, but the effect was significant only for Ala-Gly-Ser-Glu (Fig. 3*a*). When eosinophils were prepared from normal subjects, the increase in rosettes was very pronounced in the case of EA rat with Ala-Gly-Ser-Glu, Val-Gly-Ser-Glu and Val-Pro-Ser-Glu and was also significant with EA rabbit in the presence of Ala-Gly-Ser-Glu (Fig. 3*b*); however, neither Val-Gly-Ser-Glu-methyl ester nor histamine had a significant effect. The specificity of the effect for eosinophil IgG Fc receptors was investigated by examining the formation of EA rat and EA rabbit rosettes by rat and human neutrophils (Table 1). In both cases, the percentage of rosettes was higher with EA rabbit than with EA rat and none of the ECF-A tetrapeptide analogues significantly increased the extent of rosette formation. In contrast, 10^{-4} M histamine significantly suppressed the formation of EA rat rosettes with rat neutrophils without affecting the formation of other neutrophil rosettes.

Table 1 Effect of ECF-A tetrapeptides, analogues and histamine on neutrophil rosettes

	% Rat neutrophil rosettes		% Human neutrophil rosettes	
Incubation with	EA rat	EA rabbit	EA rat	EA rabbit
Medium	25.6 ± 3.2	94.0 ± 2.1	14.2 ± 6.5	99.6 ± 0.4
Ala-Gly-Ser-Glu	28.8 ± 5.4	93.3 ± 6.6	22.8 ± 3.9	100.0 ± 0
Val-Gly-Ser-Glu	34.0 ± 7.8	79.2 ± 3.9	12.5 ± 4.7	95.0 ± 5.0
Val-Pro-Ser-Glu	18.7 ± 5.3	90.2 ± 9.1	16.4 ± 6.2	98.8 ± 0.8
Val-Gly-Ser-Glu-methyl ester	23.7 ± 5.9	82.0 ± 9.4	15.2 ± 6.6	82.5 ± 12.7
Histamine	6.6 ± 1.7*	91.3 ± 3.7	10.3 ± 1.2	100.0 ± 0

Rat neutrophils were obtained from the peritoneal cavity of rats 7 h after the interaperitoneal injection of 10 ml of 10% proteose-peptone in saline. Human neutrophils were prepared from venous blood from hypereosinophilic or normal subjects according to ref. 5. Rat or human neutrophils were preincubated for 60 min at 37 °C with the various mediators at a final concentration of 10^{-4} mol l^{-1}. The values given represent the mean of five experiments ± s.e.

*The extent of suppression was significant at $P < 0.01$.

The data indicate that the human ECF-A tetrapeptides Ala[4] and Val[4], and the analogue Pro[3] markedly increase the number of Fc receptors for rat IgG on the rat eosinophil membrane. The enhancement of the expression of IgG Fc receptors on eosinophils and the augmentation of the cytotoxic capacity of rat eosinophils for IgG-coated schistosomula exhibit similar dose–response relationships, suggesting that the two phenomena are related on a fundamental cellular level. The incubation of human eosinophils with Ala-Gly-Ser-Glu and, in some cases, with Val-Pro-Ser-Glu and Val-Gly-Ser-Glu, also increased the expression of Fc receptors, whereas Val-Gly-Ser-Glu-methyl ester and histamine were inactive. Previous reports have shown no enhancement of the expression of human eosinophil IgG receptors by the ECF-A tetrapeptides, whereas complement receptors were increased by both the ECF-A tetrapeptides and histamine[8,9]. These earlier studies mainly used eosinophils from hypereosinophilic subjects and erythrocytes sensitized with rabbit antibodies. The present work clearly shows that when eosinophils from patients with hypereosinophilia were compared with normals for rosette formation with EA rabbit (Fig. 3), the patients had a higher proportion of eosinophils expressing Fc receptors, such that the increase by the tetrapeptides was only significant for one of the tetrapeptides studied. The major difference between the present report and previous studies limited to eosinophils from hypereosinophilic patients may therefore be explained by the fact that a significant increase could not be seen in the latter case because the starting percentage of Fc receptor expression was higher. It seems likely that the use of normal human eosinophils and EA-IgG rat unveiled the effects of the ECF-A tetrapeptides on the expression of eosinophil Fc receptors. Another possible explanation for the difference between the results of the two studies concerning EA-IgG rat and EA-IgG rabbit could be the preferential stimulatory role of ECF-A on surface receptors for anaphylactic antibodies which represent a more important part of rat IgG than rabbit IgG. This hypothesis is now being tested.

The large difference in the percentage of neutrophils that formed rosettes with EA rat compared with EA rabbit indicates that both the rat and human neutrophil Fc receptor seemed to possess a higher affinity for rabbit antibody than for rat anti-

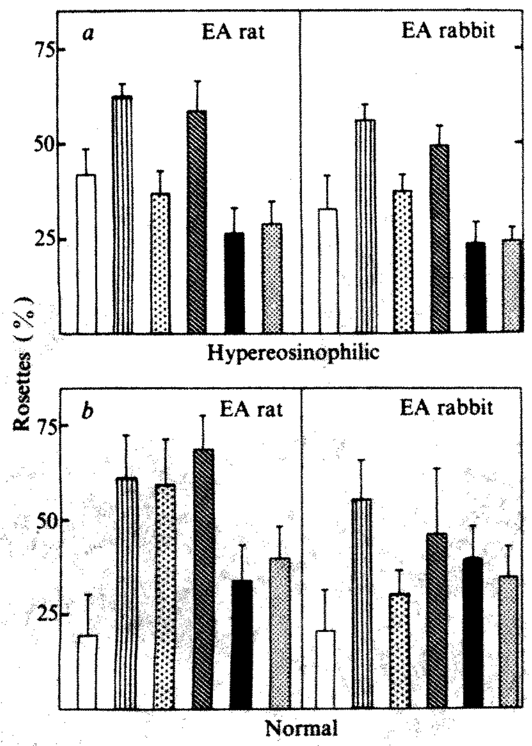


Fig. 3 The effect of the ECF-A tetrapeptides Ala-Gly-Ser-Glu and Val-Gly-Ser-Glu, the analogues Val-Pro-Ser-Glu and Val-Gly-Ser-Glu-methyl ester, and histamine on EA rat and EA rabbit rosette formation by human eosinophils. Eosinophils were prepared from venous blood of nine patients with eosinophilia of non-parasitic aetiology (*a*) and from that of four normal subjects (*b*) by centrifugation on metrizamide discontinuous gradients[5]. 80–100% pure eosinophils were obtained from both groups of individuals. EA rat and EA rabbit were prepared according to Fig. 1. □, medium; ▥, Ala-Gly-Ser-Glu; ░, Val-Gly-Ser-Glu; ▨, Val-Pro-Ser-Glu; ■, Val-Gly-Ser-Glu-methyl ester; ▩, histamine. The results are expressed as the mean of triplicate experiments (± s.e.) and were compared by a Student's *t*-test. *a*, The increase in EA rat and EA rabbit rosettes was significant with Ala-Gly-Ser-Glu ($P<0.05$). *b*, The increase in EA rat rosettes was significant with Ala-Gly-Ser-Glu, Val-Gly-Ser-Glu and Val-Pro-Ser-Glu ($P<0.05$) whereas only Ala-Gly-Ser-Glu significantly increased EA rabbit rosettes ($P<0.05$).

body. The mast cell mediators that increased the expression of eosinophil Fc receptors did not enhance either rat or human neutrophil Fc receptors, even when the percentage of Fc receptors expressed by normal neutrophils was very low (EA rat rosettes). These results therefore suggest that the effect of the tetrapeptides is specific for eosinophils.

The enhancement and inhibition of eosinophil IgG Fc receptors by mast cell-derived mediators may be a function of alterations in the rate of induction of new receptors or in the exposure of previously 'hidden' receptors, as has been suggested for complement receptors[9,10]. It has been observed that the number of erythrocytes per eosinophil is also increased, when the percentage of rosetting eosinophils is enhanced by the tetrapeptides. Further, a correlation exists between the eosinophil chemotactic activity of the tetrapeptides and their effects on the expression of IgG Fc receptors on the eosinophil membrane[11]. The parent tetrapeptides Ala^4 and Val^4 and the Pro^3 analogue exhibit the highest activity in the three assays that have been examined, namely chemotaxis, rosette formation and cytotoxicity, whereas histamine is less active or inhibitory at comparable concentrations. A relationship between surface binding sites for chemotactic factors and for the Fc fraction of IgG has been envisioned for lymphocytes[12]. Similarly, it is possible that the enhancement or inhibition of expression of eosinophil IgG Fc receptors by chemotactic factors involves an interaction between surface receptors common for chemotactic and IgG Fc binding functions, or with distinct receptors that are concurrently activated by the chemotactic factors. The capacity of the tetrapeptides of ECF-A to enhance the expression of eosinophil C3b (ref. 8) as well as IgG Fc receptors supports the latter possibility and suggests that modification of the properties of receptors may be a general membrane response of eosinophils to some chemotactic factors. The ability of defined chemotactic factors to modulate both the migration of the eosinophils and their expression of surface receptors for immunologically relevant principles indicates a more general role for chemotactic factors that should be explored for other classes of leukocytes.

A recent study reports the increased capacity of eosinophils to bind to IgG-coated particles *in vitro* when exposed to various stimuli such as soluble immune complexes or endotoxin[13]. Results of this and the present study therefore suggest possible mechanisms to account for the higher proportion of eosinophils with Fc receptors in hypereosinophilic syndromes.

The present work is thus the first demonstration that mast cell mediators like the ECF-A tetrapeptides can not only promote eosinophil recruitment but also increase IgG-mediated eosinophil cytotoxicity against *Schistosoma* targets, by enhancing the expression of eosinophil IgG Fc receptors. These results, combined with previous studies on the enhancement of complement receptors and complement-mediated eosinophil cytotoxicity[10], provide novel evidence for mast cell–eosinophil interaction in immunity against helminthic parasites.

We thank Professor M. Goudemand, Professor A. B. Tonnel and Dr L. Prin for providing human eosinophils, and J. E. Papin and S. Torres for technical assistance. This work was supported by INSERM U 167 and CNRS ERA 422.

Received 9 July; accepted 4 November 1980.

1. Goetzl, E. J. & Austen, K. F. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4123–4127 (1975).
2. Clark, R. A. F., Gallin, J. I. & Kaplan, A. P. *J. exp. Med.* **142,** 1462–1476 (1975).
3. Smith, J. A., Goetzl, E. J. & Austen, K. F. in *Peptides: Structure and Biological Function* (eds Meienhofer, J. & Gross, E.) (Pierce Chemical Co., Rockford, Illinois, in the press).
4. Capron, M., Rousseaux, J., Mazingue, C., Bazin, H. & Capron, A. *J. Immun.* **121,** 2518–2525 (1978).
5. Vadas, M., David, J. R., Butterworth, A. E., Pisani, N. T. & Siongok, T. A. *J. Immun.* **122,** 1228–1236 (1979).
6. Capron, M., Torpier, G. & Capron, A. *J. Immun.* **123,** 2220–2230 (1979).
7. Butterworth, A. E., Coombs, R. R. A., Gurner, B. W. & Wilson, A. B. *Int. Archs Allergy appl. Immun.* **51,** 368–377 (1976).
8. Anwar, A. R. E. & Kay, A. B. *Nature* **269,** 522–524 (1977).
9. Anwar, A. R. E. & Kay, A. B. *J. Immun.* **121,** 1245–1250 (1978).
10. Kay, A. B. *J. Allergy clin. Immun.* **64,** 90–104 (1979).
11. Capron, M., Goetzl, E. J., Austen, K. F. & Capron, A. *Prog. Immun.* (in the press).
12. Shields, J. M. & Wilkinson, P. C. *Clin. exp. Immun.* **38,** 598–608 (1979).
13. Tai, P. C. & Spry, C. J. F. *Clin. exp. Immun.* **40,** 206–219 (1980).
14. Lynch, S. M., Austen, K. F. & Wasserman, S. I. *J. Immun.* **121,** 1394–1399 (1978).

Synaptic localization of kainic acid binding sites

A. C. Foster, E. E. Mena, D. T. Monaghan & C. W. Cotman

Department of Psychobiology, University of California, Irvine, Irvine, California 92717

The heterocyclic compound kainic acid (KA) is a potent excitant when applied to mammalian neurones[1]. Lesions caused by injections of KA into the rat striatum and hippocampus cause similar patterns of damage to those seen in Huntington's chorea and status epilepticus, respectively[2,3]. Although it was originally thought to be a glutamate agonist[4], it is now clear that KA does not act on the majority of the receptors for glutamate[5–7], and in fact seems to act on a class of receptors which are distinct from those which mediate responses to other excitatory amino acids[8,9]. The potent and selective neurotoxic effects of this compound[10] may be mediated by these same receptors. At present, the relative distribution of junctional and extrajunctional (non-synaptic) receptors is unknown and resolution of this issue would provide important insights into the action of KA on the central nervous system (CNS). We show here that KA binding sites are greatly enriched in isolated synaptic junctions from rat brain and, using an *in vitro* autoradiographic technique, we have found that these binding sites are concentrated specifically in terminal fields where KA acts as a potent neurotoxin.

Synaptic junctions (SJs) and other subcellular fractions were prepared by the method of Cotman and Taylor[11], and assayed for ^{3}H-KA binding[12,13]. Table 1 shows the relative values for specific binding (the ^{3}H-KA binding which can be displaced by 100 μM unlabelled KA) in the isolated fractions, which are expressed as a value relative to the whole particulate fraction (=1). Membranes from crude mitochondrial pellet (P_2) were not enriched in KA binding sites relative to the whole particulate fraction. However, on subfractionation of P_2, the highest specific binding values were found in the synaptic plasma membranes (SPMs), in agreement with previous results[13]. The P_1, P_3, myelin enriched and mitochondria enriched fractions also had specific KA binding sites. This may be due to the presence of synaptic plasma membranes in these fractions, and indeed both purified myelin and mitochondria have been found to contain lower levels of high-affinity binding sites for ^{3}H-KA[13].

Specific KA binding sites in SJs were enriched 22 times over those measured in the whole brain particulate fraction, and 10 times over those in SPMs. This enrichment in SJ fractions is greater than that shown previously for any other receptor or receptor-related function. The absolute amount of binding to SJs, averaged over four separate preparations and assayed at a concentration of 40 nM, was 1,819 ± 378 fmol per mg protein. The majority of specific KA binding sites in the SPMs could be recovered in SJs (73 ± 14% of binding sites; 9.4 ± 0.6% protein; $n = 4$); which suggests that most of the binding sites for KA in SPMs are located at the junctional area. Similarly, 80% of the binding sites in SPMs for L-glutamate and L-aspartate are recovered in SJs; quinuclidinyl benzoate binding sites are not detectable in SJs prepared by this method.

Two components of KA binding were apparent in SJs (assayed over a ^{3}H-KA concentration range of 1–100 nM). The mean values from two determinations measured for the high-affinity component were $K_D = 4.95$ nM and $B_{max} = 1{,}349$ fmol per mg protein, and for the low-affinity component 32.6 nM and 3,257 fmol per mg protein, respectively. These values for the K_Ds agree with those previously reported in crude membrane preparations[14], and indicate that the increase of KA binding seen in SJs is a true enrichment of sites and is not due to an increase in affinity caused, for example, by removal of an endogenous inhibitor.

If KA binding sites are predominantly synaptic, they might be localized to discrete terminal fields. KA binding sites can be viewed directly in rat brain slices by an *in vitro* autoradiographic technique[15]. Using this method, KA binding sites were seen to be particularly enriched in discrete parts of the hippocampus, pyriform cortex, cerebellum and striatum. In the hippocampus a small band of silver grains was associated with the inner part of the molecular layer of the dentate gyrus, corresponding to the terminal field of the commissural/associational fibres. However, the greatest density of grains was seen over the stratum incidium in the CA3 region corresponding to the terminal zone of the mossy fibre input to the CA3 pyramidal cells (Fig. 1). The KA binding sites labelled by this technique were apparently identical to those revealed in membrane binding assays, as this pattern of binding could be completely displaced with an excess of unlabelled KA, and binding to these slices was saturable with a K_D value in the range found for SJs. The pattern of KA binding sites we have observed in the hippocampus is particularly interesting in view of the previous reports that the CA3 pyramidal cells of the hippocampus are extremely sensitive to the neurotoxic action of KA[16], particularly at the mossy fibre terminal zone[17]. Other brain areas also showed KA sites but in general these were of lower concentration than the CA3 field (manuscript in preparation).

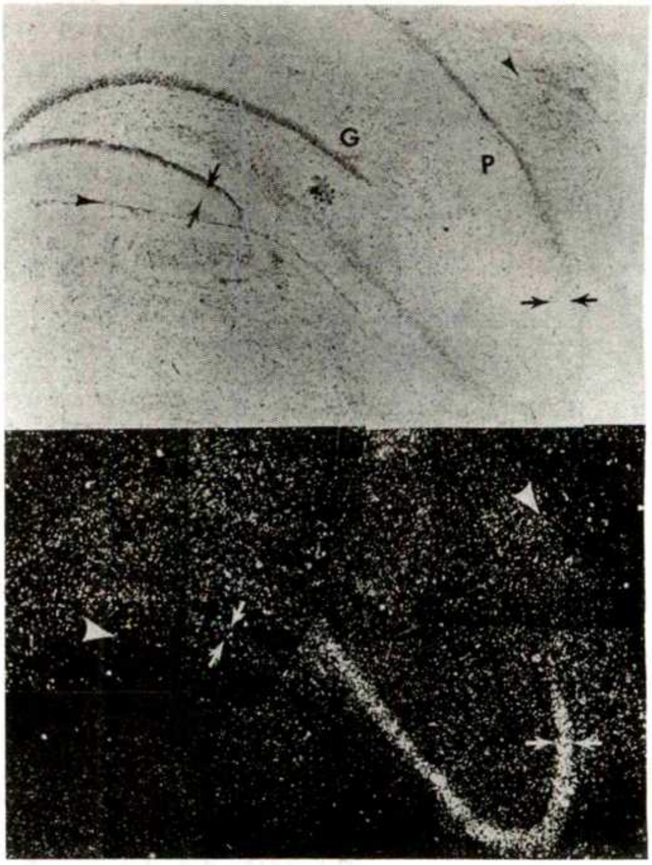


Fig. 1 Distribution of KA binding sites in the rat hippocampus. Brain slices (15 μm) from adult Sprague–Dawley rats were sliced on a cryostat, mounted unfixed on glass microscope slides and processed for the localization of binding sites using a modification of the method of Young and Kuhar[15]. ^{3}H-KA was used at a final concentration of 100 nM, and incubated at 0 °C for 30 min. Bottom: dark field illumination, small arrows indicate the areas of high silver grain density. Top: cresyl violet-stained tissue section; small arrows indicate identical area in dark field possessing high silver grain density. G, granule cells; P, pyramidal cells.

Table 1 Specific binding of ^{3}H-KA to subcellular fractions of rat brain

Fraction	Specific ^{3}H-KA binding relative to OH (=1)	n
Whole particulate (OH)	1	4
Crude mitochondrial pellet (P_2)	0.73 ± 0.34	4
Microsomes (P_3)	0.56, 0.25	2
Crude myelin	0.05, 0.13	2
Synaptic plasma membrane (SPM)	2.08 ± 0.41	4
Mitochondria	0.95 ± 0.21	3
Synaptic junction (SJ)	22.37 ± 5.79	4

Subcellular fractions were prepared from whole rat forebrain as described by Cotman and Taylor[11]. All fractions were washed in 0.2 mM HEPES, 0.05 mM $CaCl_2$ buffer (pH 7.4) and pelleted at 80,000g for 20 min. Binding assays were done in triplicate essentially as previously described[13]. Crude myelin was taken as the fraction floating on the 0.8 M sucrose layer. The mitochondria fraction was the 1.2 M pellet.

The results presented here, therefore, suggest a discrete localization of KA binding sites in the mammalian CNS. At a subcellular level, SJs contribute the majority of the 'synaptic population' of KA binding sites, and in the hippocampus these sites are present in high density in close apposition to the mossy fibre–CA3 pyramidal cell synapse. Neither glutamate nor aspartate seem to be transmitters at this synapse[18]. Specific and localized synaptic receptors imply that appropriate natural ligands exist. Our data, together with evidence that KA does not act on glutamate sites and that KA receptors are a discrete population[8,9], suggest that endogenous ligands for KA receptors exist. Such a compound would be a mediator or modulator of synaptic transmission at sites where we have shown KA receptors to exist. Furthermore, the identification of such compounds may provide new insights into the aetiology of Huntington's disease and epilepsy.

This work was supported by NIH and NIMH grants. A.C.F. holds a NATO postdoctoral fellowship and E.E.M. a NIH postdoctoral fellowship. We thank Mrs D. Franks for assistance.

Received 2 September; accepted 27 October 1980.

1. Shinozaki, H. & Konishi, S. *Brain Res.* **24,** 368–371 (1970).
2. Coyle, J. T., McGeer, E. G., McGeer, P. L. & Schwarcz, R. in *Kainic Acid as a Tool in Neurobiology* (eds McGeer, E. G., Olney, J. W. & McGeer, P. L.) 139–160 (Raven, New York, 1978).
3. Nadler, J. V., Perry, B. W. & Cotman, C. W. in *Kainic Acid as a Tool in Neurobiology* (eds McGeer, E. G., Olney J. W. & McGeer, P. L.) 219–238 (Raven, New York, 1978).
4. Johnston, G. A. R., Curtis, D. R., Davies, J. & McCulloch, R. M. *Nature* **248,** 804–805 (1974).
5. Hall, J. G., Hicks, T. P. & McLennan, H. *Neurosci. Lett.* **8,** 171–175 (1978).
6. Foster, A. C. & Roberts, P. J. *Neurochemistry* **31,** 1467–1477 (1978).
7. London, E. D. & Coyle, J. T. *Eur. J. Pharmac.* **56,** 287–290 (1979).
8. Davies, J., Evans, R. H., Francis, A. A., Jones, A. W. & Watkins, J. C. in *EMBO Workshop on Drug Receptors in the Central Nervous System* (ed. Balaban, M.) (Int. Sci. Services, Jerusalem, 1980).
9. McLennan, H. & Lodge, D. *Brain Res.* **169,** 83–90 (1979).
10. McGeer, E. G., Olney, J. W. & McGeer, P. L. (eds) *Kainic Acid as a Tool in Neurobiology* (Raven, New York, 1978).
11. Cotman, C. W. & Taylor, D. *J. Cell Biol.* **55,** 696–711 (1972).
12. Simon, J. R., Contrera, J. F. & Kuhar, M. J. *J. Neurochem.* **26,** 141–147 (1976).
13. Nieto-Sampedro, M., Shelton, D. & Cotman, C. W. *Neurochem. Res.* **5,** 591–604 (1980).
14. London, E. D. & Coyle, J. T. *Molec. Pharmac.* **15,** 492–505 (1979).
15. Young, W. S. & Kuhar, M. J. *Brain Res.* **179,** 255–270 (1979).
16. Nadler, J. V., White, W. F., Vaca, K. W. & Cotman, C. W. *Nature* **271,** 676–677 (1978).
17. Olney, J. W., Fuller, T. & DeGubareff, T. *Brain Res.* **176,** 91–100 (1979).
18. Nadler, J. V., White, W. F., Vaca, K. W., Perry, B. W. & Cotman, C. W. *J. Neurochem.* **31,** 147–155 (1978).

Corticotropin regulates trans-synaptically the activity of septo-hippocampal cholinergic neurones

Lawrence J. Botticelli* & Richard J. Wurtman

Laboratory of Neuroendocrine Regulation, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The activity of septo-hippocampal neurones is affected by the action on cholinergic perikarya in the septum of a variety of putative neurotransmitters[1–4], including substance P[5,6] and β-endorphin[7,8]. (The latter is released in the septal region from neurones which originate in the medial basal hypothalamus.) It has also been reported that two other neuropeptides, corticotropin ($ACTH_{1-24}$) and α-melanotropin (α-MSH), affect acetylcholine turnover in septo-hippocampal neurones in a manner that is not blocked by transection of the afferents to the hippocampus, from which it has been inferred that the neurotransmitters act directly on the hippocampus[9]. We now describe experiments with corticotropin which show that the effect is rather the influence on septo-hippocampal cholinergic neurones of peptidergic neurones within the septum.

Male Sprague–Dawley rats (150–200 g, Charles River Breeding Laboratories) were housed in groups of three to five at an ambient temperature of 23–24 °C for 5–7 days before experimentation. Animals had free access to food (Purina Rat chow; Ralston-Purina) and water, and were exposed to light (Vita-Lite, 300 μW cm^{-2}; Duro Test Corporation) daily between 0700 and 1900 h.

A polyethylene tube was implanted in the lateral ventricle (2 mm lateral, 1 mm posterior to bregma and 3 mm ventral to the dural surface) of each rat 3 days before experimentation. Cannulae were kept in position with dental cement. For intraseptal injections, guide cannulae were implanted stereotaxically into the septum (coordinates AP 8.5, ML 0.0, DV 0.0, 5.0 mm ventral to the dural surface) 5 days before experimentation. For intrahippocampal injections, cannulae were placed in the dorsal hippocampus (coordinates AP 4.0, ML 2.0, DV 2.0) 3 days before experimentation. Cannulae placements were verified histologically by staining a 1 : 3 or 1 : 5 series of 40-μm frozen frontal cryostat sections with thionin. For peptide co-administration studies, rats were equipped with both intraventricular and intraseptal cannulae.

Table 1 Effects of peptides on hippocampal acetylcholine (ACh) content

Treatment	Hippocampal ACh (pmol per mg)
Vehicle	40.0 ± 2.3 (n = 6)
β-lipotropin	37.5 ± 1.7 (n = 5)
β-endorphin	51.4 ± 2.7* (n = 5)
Corticotropin	25.5 ± 1.5* (n = 5)
$ACTH_{4-10}$	38.7 ± 2.9 (n = 5)
α-MSH	39.9 ± 2.2 (n = 5)

Rats were killed 30 min after intraventricular administration of each peptide. Values are expressed as means ± s.e.m. Data were analysed by analysis of variance (corrected for unequal sample size), followed by *a posteriori* multiple comparisons of within-group means using Scheffe's method[11].

* $P < 0.001$ differs from vehicle-treated animals (*a posteriori*).

Peptides dissolved in pyrogen-free water were microinjected into the lateral ventricle (10 μg, into the septum (1.0 μg in 1.0 μl given over 5 min), or the dorsal hippocampus (1.0 μg in 1.0 μl given over 3 min). At 30 min after peptide micro-administration the animals were killed by a beam of microwave irradiation directed at the skull. After decapitation, brains were removed from the skull and specimens of cerebral cortex, striatum and hippocampus (dorsal) dissected by hand; their acetylcholine and choline content was determined radioenzymatically using [γ-^{32}P]ATP (specific activity 10–30 Ci nmol^{-1}; Amersham) as labelled substrate[10]. Significance of treatment-induced differences in regional brain acetylcholine or choline content was determined by analysis of variance (Wang 2200 series software; Wang Laboratories), followed by *a posteriori* multiple comparisons of within-group means using Scheffe's method[11].

Acetylcholine content in cortex, striatum and hippocampus was unchanged after intraventricular administration of β-LPH (Table 1). As described previously[8], intraventricular β-endorphin significantly increased acetylcholine content in the hippocampus (38%; $P < 0.001$) (Table 1) but not elsewhere. Corticotropin had an opposite effect on hippocampal acetylcholine concentrations, causing them to fall by about 30% ($P < 0.001$) (Table 1), but like β-endorphin, failed to affect acetylcholine levels in the two other brain regions examined. Neither $ACTH_{4-10}$ nor α-MSH affected acetylcholine concentrations in any brain region examined (Table 1). None of the peptides affected choline levels in any region (data not shown).

When smaller doses (1 μg) of β-endorphin or corticotropin were microinjected directly into the septum, hippocampal acetylcholine levels changed in the same direction as after its

Table 2 Effects of β-endorphin and/or corticotropin on hippocampal acetylcholine

Treatment	Hippocampal ACh (pmol per mg)
Vehicle	37.7 ± 1.2 (n = 4)
β-endorphin (i.s.t.)	58.4 ± 3.0* (n = 5)
Corticotropin (i.s.t.)	21.6 ± 2.4* (n = 5)
β-endorphin + corticotropin (i.s.t.)	42.0 ± 3.1† (n = 6)
β-endorphin (i.v.t.) + corticotropin (i.s.t.)	39.9 ± 1.3† (n = 6)
β-endorphin (i.s.t.) + corticotropin (i.v.t)	37.9 ± 2.6† (n = 6)

Rats were killed 30 min after injection of β-endorphin and corticotropin [either both intraseptally (i.s.t.) (0.5 μg each) or one intraseptally (1.0 μg) and the other intraventricularly (i.v.t.) (10 μg)]. Values are expressed as means ± s.e.m. Data were analysed by analysis of variance (corrected for unequal sample size) followed by *a posteriori* multiple comparisons of within-group means using Scheffe's method[11].

* $P < 0.001$ differs from vehicle-treated animals (*a posteriori*).

† $P < 0.001$ differs from peptide-treated animals (*a posteriori*).

---

* Present address: Addiction Research Foundation, 701 Welch Road, Palo Alto, California 94304.

0028-0836/81/020075—02$01.00

intraventricular injection but to a greater extent: β-endorphin administration increased acetylcholine concentrations by 56% ($P < 0.001$), whereas corticotropin decreased hippocampal acetylcholine by 44% ($P < 0.001$) (Table 2). Co-administration of the neuropeptides, either both intraseptally (0.5 μg each), or one intraseptally (1.0 μg) and the other intraventricularly (10 μg), did not significantly change hippocampal acetylcholine levels (Table 2). Direct administration of β-endorphin or corticotropin into the dorsal hippocampus (1.0 μg) also failed to affect hippocampal acetylcholine (Table 3).

Corticotropin and β-endorphin, which may share a common precursor[12], have similar brain distributions; they are concentrated within a major cell group in the periarcuate hypothalamus, projecting to the anterior hypothalamus, pons, the periaqueductal grey and the septal region[13-15]. Nerve fibres immunoreactive for both corticotropin and β-endorphin surround the third ventricle, extend throughout the diencephalon and portions of the limbic brain, and descend into the central grey. No corticotropin or endorphin immunoreactivity has been detected in the hippocampus, cerebrum or cerebellar cortices. Corticotropin and β-endorphin can sometimes block each other's effects, or produce effects opposite to one another[16-18], which suggests that these peptidergic neurones may participate reciprocally in common neuronal systems.

Table 3 Effects of intrahippocampal β-endorphin or corticotropin on hippocampal acetylcholine

Treatment	Hippocampal ACh (pmol per mg)
Vehicle	39.6 ± 2.1 ($n = 5$)
β-endorphin	43.6 ± 1.5 ($n = 5$)
Corticotropin	41.9 ± 2.8 ($n = 5$)

Rats were killed 30 min after injection of peptides into the dorsal hippocampus (1.0 μg per 1.0 μl for 3.0 min). Values are expressed as means ± s.e.m.

The activity of septo-hippocampal neurones is controlled partly by axosomatic and axodendritic synapses in the septum itself, and partly by hippocampal axoaxonic synapses. Intraseptally injected β-endorphin produces a naltrexone-reversible decrease in the turnover of hippocampal acetylcholine[7], suggesting that endorphin interacts with opiate receptors in the septal region. In contrast, corticotropin has been reported to increase the turnover of this neurotransmitter in Ammon's horn, even after transection of the cingulum, entorhinal area and fimbria[9]. Such observations have been interpreted as suggesting a direct action of these neuropeptides on receptors in the hippocampal formation. Our observations that intraseptal corticotropin decreases hippocampal acetylcholine, that intrahippocampal corticotropin fails to change the content of hippocampal acetylcholine and that co-administered β-endorphin and corticotropin block each other's effects on hippocampal acetylcholine, imply that corticotropin-containing peptidergic neurones do act within the septum, and thereby influence the physiological activity of septo-hippocampal cholinergic neurones.

These studies were supported in part by research grants awarded to R.J.W. from the National Institute of Mental Health (MH 28783), the Ford Foundation and the Wallace Genetic Foundation. L.J.B. was a predoctoral fellow of the National Institute on Drug Abuse (DA 05089). β-Lipotropin was supplied by Dr C. H. Li, β-endorphin by the Psychopharmacology Research Branch, National Institute of Mental Health, corticotropin, $ACTH_{4-10}$ and α-MSH were gifts of Organon Pharmaceuticals, and all other chemicals used were purchased from either Eastman Organic or Aldrich.

Received 15 September; accepted 24 October 1980.

1. McLennan, H. & Miller, J. J. *J. Physiol., Lond.* **237,** 607–624, 624–633 (1974).
2. DeFrance, J. F., Yoshihara, H., McCrea, R. A. & Kitai, S. T. *Expl Neurol.* **48,** 502–523 (1975).
3. Assaf, S. Y. & Miller, J. J. *Brain Res.* **129,** 353–360 (1977).
4. Moore, R. Y. *J. comp. Neurol.* **177,** 665–684 (1978).
5. Cuello, A. C. & Kanazawa, I. *J. comp. Neurol.* **178,** 129–156 (1978).
6. Malthe-Sorenssen, D., Cheney, D. L. & Costa, E. *J. Pharmac. exp. Ther.* **206,** 21–28 (1978).
7. Moroni, F., Cheney, D. L. & Costa, E. *Naunyn-Schmiedebergs Archs Pharmak.* **299,** 149–153 (1977).
8. Botticelli, L. J. & Wurtman, R. J. *Life Sci.* **24,** 1799–1804 (1979).
9. Wood, P. L., Cheney, D. L. & Costa, E. *J. Pharmac. exp. Ther.* **209,** 97–103 (1979).
10. Goldberg, A. M. & McCaman, R. E. *J. Neurochem.* **20,** 1–8 (1973).
11. Winer, B. *Statistical Principles in Experimental Design*, 205–210 (McGraw-Hill, New York, 1971).
12. Eipper, B. A. & Mains, R. E. *Endocr. Rev.* **1,** 1–27 (1980).
13. Bloom, F. E. *et al.* in *The Endorphins* (eds Costa, E. & Trabucchi, M.) 89–109 (Raven, New York, 1900).
14. Watson, S. J., Akil, H., Richards, C. W. III & Barchas, J. D. *Nature* **275,** 226–228 (1978).
15. Watson, S. J., Richards, C. W. III & Barchas, J. D. *Science* **200,** 1180–1182 (1978).
16. Jacquet, Y. F. *Science* **201,** 1032–1034 (1978).
17. Simantov, R. *Nature* **280,** 684–685 (1979).
18. Botticelli, L. J. & Wurtman, R. J. in *Endogenous and Exogenous Opiate Agonists and Antagonists* (ed. Leong Way, E.) 187–190 (Pergamon, New York, 1980).

Serotonin stimulates phosphorylation of Protein I in the facial motor nucleus of rat brain

Annette C. Dolphin* & Paul Greengard†

Department of Pharmacology, Yale University School of Medicine,
New Haven, Connecticut 06510

Protein I is one of the best candidates for a neuronal protein whose phosphorylation may have a functional role in synaptic activity[1]. It is a substrate for both cyclic AMP-dependent[2] and Ca^{2+}-dependent[3] protein kinases, and these kinases show differential specificity for its multiple phosphorylation sites[4]. Protein I is found exclusively in the central[5] and peripheral[6] nervous systems, and immunohistochemical[5] and subcellular fractionation[7] studies suggest an association primarily with synaptic vesicles. Using slices of rat cerebral cortex incubated *in vitro*, Protein I was phosphorylated both by agents which increase intracellular cyclic AMP and by agents causing Ca^{2+} influx, although not by any putative neurotransmitters or neuromodulators[8]. We have now examined the facial motor nucleus and report here that serotonin produces a phosphorylation of Protein I when incubated with facial nucleus slices. Demonstration of a neurotransmitter-dependent alteration in the state of phosphorylation of a synapse-specific protein may be due to the relatively simple neuronal circuitry within the facial motor nucleus.

The facial motor nucleus contains a single class of neuronal cell bodies and no interneurones[9,10]. Most of the afferent fibres are excitatory onto the facial motoneurones[10] and stimulation of motoneurone firing is mimicked by the iontophoresis of glutamate[10]. The facial nucleus also receives a facilitatory serotonergic input from the raphe magnus[10-12], which represents about 2% of the terminals in the nucleus[13].

Slices prepared from excised rat facial nuclei were preincubated for 1 h in Krebs–Ringer bicarbonate buffer (KRB). The slices were then incubated for 1–20 min in the presence of serotonin or other agents. The reaction was terminated by homogenizing the tissue in zinc acetate (5 mM), which inhibits phosphatase activity[8]. Protein I was quantitatively extracted and subjected to 'back-phosphorylation' by the catalytic subunit of cyclic AMP-dependent protein kinase, using [γ-^{32}P]ATP as substrate. This method of back-phosphorylation gives a measure of the amount of dephospho-Protein I remaining in the slice at the end of the incubation period, from which one can calculate the amount of Protein I that was converted to the phosphorylated form during incubation with the various agents[8] (see Figs 1 and 2 legends).

In resting conditions, after preincubation of the slices for 1 h, Protein I seemed to be almost entirely in the dephosphorylated

---

* Present address: Division of Neurophysiology and Neuropharmacology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.
† To whom reprint requests should be addressed.

0028-0836/81/020076—04$01.00

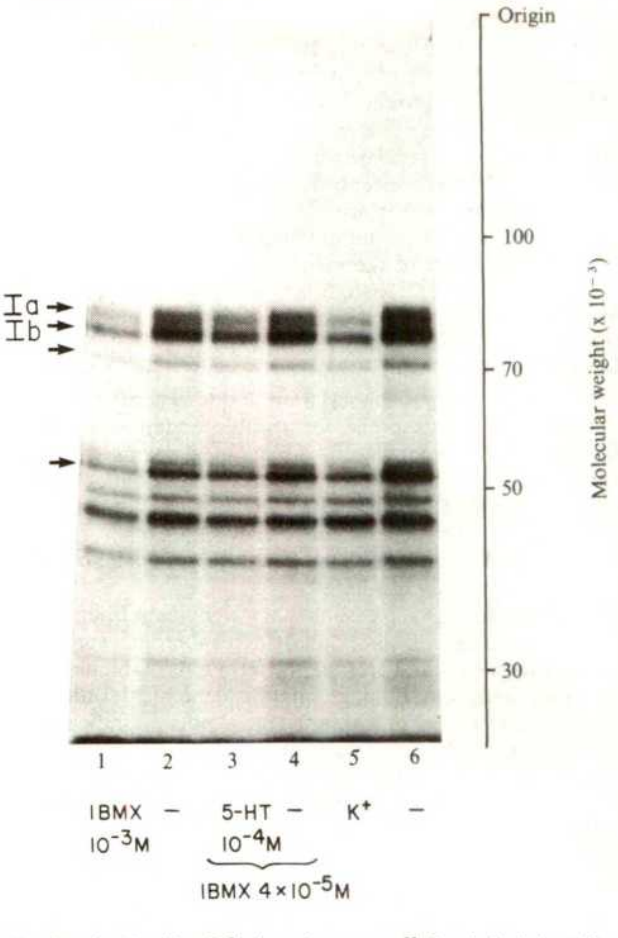


Fig. 1 Autoradiogram of an SDS-polyacrylamide gel showing the effect of IBMX, of serotonin (5-HT) and of high K^+ on protein phosphorylation in slices of rat facial nucleus. Male Sprague–Dawley rats (150–250 g) were killed by decapitation. The brain was removed and a transverse slice, 1.2 mm thick, was taken through the brain stem immediately posterior to the trapezoid body. The regions containing the facial nuclei were then dissected manually as cubes taken from the ventral edge of the slice with the pyramidal tracts as the medial borders. Each excised nucleus was quartered using a sagittal and a horizontal cut. The four slices from a single facial nucleus were transferred together to a glass homogenizer containing 10 ml of KRB, *p*H 7.4, at 35 °C (composition in mM: Na^+, 158; K^+, 3; Ca^{2+}, 1.5; Mg^{2+}, 1.3; PO_4^{3-}, 1.3; HCO_3^-, 26; Cl^-, 136; ascorbate, 1; D-glucose, 20). The medium was continuously gassed with H_2O-saturated 95% O_2–5% CO_2 throughout the preincubation and incubation periods. Slices were preincubated for 1 h with one change of medium. In all cases, the effect of the agent to be tested was determined by incubating it for 1 or 10 min with one of a pair of facial nuclei, the other of which was used as a control. Substances to be tested were added in KRB such that the final incubation volume was 10 ml. The incubations were terminated by aspirating the KRB, adding 5 ml of ice-cold Zn-acetate (5 mM) and homogenizing with a Teflon pestle. The amount of Protein I present in the dephospho-form at the termination of the incubation was determined by back-phosphorylation as follows: the homogenate was centrifuged at 4,000*g* for 15 min and the pellet resuspended in 1 ml of 10 mM Na-citrate-phosphate buffer, *p*H 2.8, at 4 °C. Triplicate 100-μl aliquots of this suspension were taken for protein assay using a method with which monoamines did not interfere[31]. The suspension was centrifuged at 23,000*g* for 15 min and 500 μl of the supernatant was adjusted to *p*H 6 with 0.5 M Na_2HPO_4 and centrifuged at 15,000*g* for 15 min. All *p*H 6 supernatants from one experiment were then diluted to equal protein concentration with 10 mM Na-citrate-phosphate buffer (*p*H 6). Duplicate or triplicate 60-μl aliquots were used in the back-phosphorylation assay containing (in 100-μl final volume): 50 mM HEPES (*p*H 7.4), 10 mM Mg^{2+}, 1 mM EGTA, 1 mM EDTA, 3 μM [γ-^{32}P]ATP (specific activity 50–100 c.p.m. $fmol^{-1}$) and a low concentration (7.5 nM) of the purified catalytic subunit (C) of cyclic AMP-dependent protein kinase, prepared according to Beavo *et al.*[32]. The back-phosphorylation assay was carried out for 30 min at 30 °C and samples were subjected to SDS-polyacrylamide gel electrophoresis and autoradiography as previously described[8]. In control experiments, it was established that the back-phosphorylation assay conditions used resulted in a maximal incorporation of ^{32}P into phospho-peptide 1 of Protein I, with negligible phosphorylation of other phosphorylation sites on Protein I (ref. 33, and A.C.D., unpublished observations). Phospho-peptide 1 is the only tryptic phosphopeptide derived from Protein I which shows a high affinity for phosphorylation by cyclic AMP-dependent protein kinase, although it is also phosphorylated in a Ca^{2+}-dependent manner[33]. Protein I is indicated by two arrows. It consists of two subspecies Ia and Ib of M_r 86,000 and 80,000. Lanes 1 and 2 incubations were with or without 10^{-3} M IBMX for 10 min; lanes 3 and 4 incubations were with or without 10^{-4} M serotonin for 10 min, in the presence of a low concentration of IBMX (4×10^{-5} M) added to both incubation tubes 2 min before the start of incubation; lanes 5 and 6 incubations were in the presence or absence of 60 mM K^+ (K^+ being substituted isotonically for Na^+) for 1 min.

form. A 10-min incubation with serotonin (10^{-4} M) increased the phosphorylation of Protein I in facial nucleus slices (Fig. 1, lane 3 compared with 4), as did the phosphodiesterase (PDE) inhibitor isobutyl methylxanthine (IBMX) (10^{-3} M; Fig. 1, lane 1 compared with 2). Each of these agents raises cyclic AMP levels in facial nucleus slices[14]. Several other proteins also showed increases in their state of phosphorylation in response to incubation of slices with serotonin and IBMX, indicating that they are also endogenous substrates for cyclic AMP-dependent protein kinase. The latter demonstrations are reasonable in view of the methodology used for the back-phosphorylation assay, in which the low concentration of catalytic subunit used selectively phosphorylated only those substrate proteins with a high affinity for this enzyme. As neurotransmitters thought to act through cyclic AMP can cause numerous biological responses in their target cells (for example, alterations in membrane excitability, in intermediary metabolism and in protein synthesis), many different phosphoproteins are probably involved in these responses[15,16]. Indeed, several endogenous substrates for cyclic AMP-dependent protein kinase have already been detected in isolated neuronal membrane preparations[2,17–20].

Incubation of facial nucleus slices with a high concentration of K^+ (60 mM) also increased the phosphorylation of Protein I and several other phosphoproteins (Fig. 1, lane 5 compared with 6). That high K^+ altered the state of phosphorylation of several of the substrate proteins affected by serotonin suggests that regulation of phosphorylation sites on proteins by both cyclic AMP-and Ca^{2+}-dependent protein kinases may not be unusual. In this regard, note that two of the affected proteins, indicated by arrows in Fig. 1, seem to be the same as the 74,000 and 55,000 molecular weight (M_r) proteins whose phosphorylation was affected by 8-bromo-cyclic AMP, high K^+ and veratridine in slices of rat cerebral cortex[8].

The formation of phospho-Protein I from dephospho-Protein I, as a function of serotonin concentration, is shown in Fig. 2. The stimulation of phosphorylation caused by serotonin was potentiated by IBMX (Fig. 2) at a concentration (4×10^{-5} M) which alone caused little alteration in the state of Protein I phosphorylation. In the presence of this concentration of IBMX, the maximal effect of a 10-min incubation with serotonin (10^{-3} M) was to convert $30.1\pm3.7\%$ ($n=9$) of Protein I from the dephospho- to the phospho-form, and the EC_{50} of serotonin was 4×10^{-6} M. The time course of Protein I phosphorylation by serotonin (10^{-3} M) in the presence of IBMX (4×10^{-5} M) was investigated and found to be similar to that observed for 10^{-3} M IBMX in cortical slices[8]. The conversion of dephospho-Protein I to phospho-Protein I was half-maximal after about 2 min, maximal after 5 min and remained constant for at least another 15 min.

That the maximal phosphorylation of Protein I in response to serotonin was not greater than 30% may be because, to excise the facial nucleus intact, it was necessary to take some tissue from adjacent regions which are not innervated by serotonergic terminals. Moreover, in all tissue-slice systems studied so far (for review, see ref. 15), 30–50% of Protein I has been observed to be refractory to phosphorylation in steady-state conditions by 8-bromo-cyclic AMP, IBMX, K^+ or veratridine, either alone or in combination. In the present study, incubation with 10^{-3} M IBMX for 10 min produced a $50.2\pm3.4\%$ ($n=8$) phosphorylation of Protein I, and incubation of facial nucleus slices for 1 min with 60 mM K^+-KRB produced a $48.5\pm3.6\%$ ($n=7$) phosphorylation of Protein I. The nature and subcellular localization of the refractory Protein I is being investigated.

That serotonin affected Protein I phosphorylation by interaction with specific receptors was indicated by the ability of mianserin, an antagonist at central serotonin receptors[21–23], to block the phosphorylation of Protein I induced by 10^{-4} M serotonin (Fig. 3). However, 10^{-5} M mianserin also caused a slight phosphorylation of Protein I when incubated with slices in the absence of serotonin. This and other classical serotonin antagonists have been found to be partial agonists of serotonin sensitive adenylate cyclase in brain, in that they stimulate adenylate cyclase at concentrations of 10^{-5} M and above[23].

In addition to the serotonergic input to the facial nucleus, there is a noradrenergic input which has been estimated to be about one-fifth that of the serotonin innervation of the facial nucleus in the rat[10,24]. Also, functional α-adrenergic receptors have been demonstrated on the facial motor neurones[10]. However, (−)noradrenaline (10^{-3} or 10^{-4} M) with or without 4×10^{-5} M IBMX, and (−)phenylephrine (10^{-3} M), had no affect on Protein I phosphorylation in this system.

In view of the widespread occurrence of adenosine receptors associated with adenylate cyclase in nervous tissue, we examined the effect of the adenosine analogue 2-chloroadenosine (10^{-3} M) on the phosphorylation of Protein I in facial nucleus slices. This agent converted 33.3±6.5% ($n=4$) of Protein I to the phospho-form. The effects of adenosine analogues are reported elsewhere[25].

We conclude for two reasons that a significant fraction of the Protein I that is phosphorylated in response to serotonin is probably located in the presynaptic terminals that innervate the facial motoneurones. First, immunohistochemical studies of the distribution of Protein I within the facial nucleus indicate that it is present in most and possibly all nerve terminals, whereas it was not detectable in the motoneurone somata (P. De Camilli and A. C. D., unpublished results). Second, serotonin has been found to stimulate selectively the phosphorylation of phosphopeptide 1 of Protein I in synaptosomes prepared from the rat facial nucleus: these results were obtained by preincubating synaptosomes with ^{32}P-inorganic phosphate to prelabel the presynaptic ATP pool, and then measuring the incorporation of ^{32}P into Protein I in response to incubation of the synaptosomes with serotonin[26].

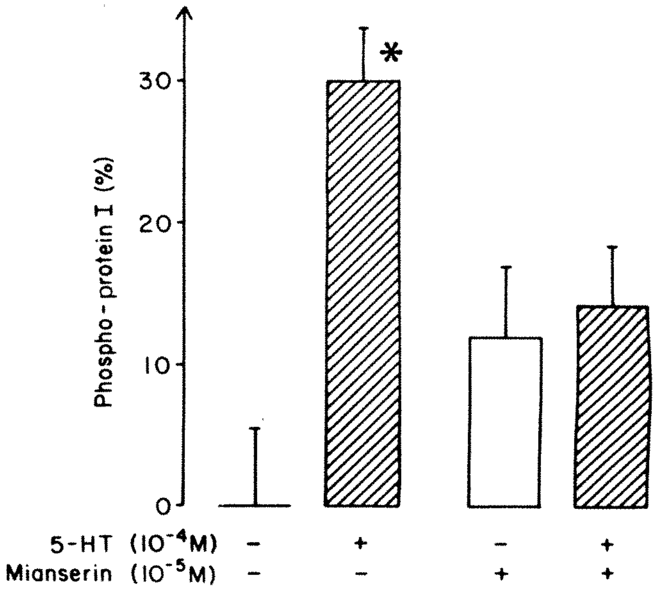


Fig. 3 The effect of the serotonin antagonist mianserin on the phosphorylation of Protein I in rat facial nucleus slices. One of each pair of facial nucleus slices was incubated with, and one without, serotonin (5-HT, 10^{-4} M) for 10 min. Mianserin (10^{-5} M), where included, was added to each pair 15 min before the end of the incubation. IBMX (4×10^{-5} M) was present in all incubation tubes and was added 12 min before the end of the incubation. The amount of dephospho- and phospho-Protein I was determined as described in Fig. 2 legend and was calculated with reference to slices incubated in the absence of both serotonin and mianserin. The basal level of phosphate incorporation into Protein I in the absence of serotonin and mianserin was 6.7±0.4 pmol per mg Zn acetate-precipitable protein. Data represent the mean±s.e.m. for five samples, each assayed in duplicate.*, $P<0.01$ when the effect of serotonin is compared with basal level (Student's t-test).

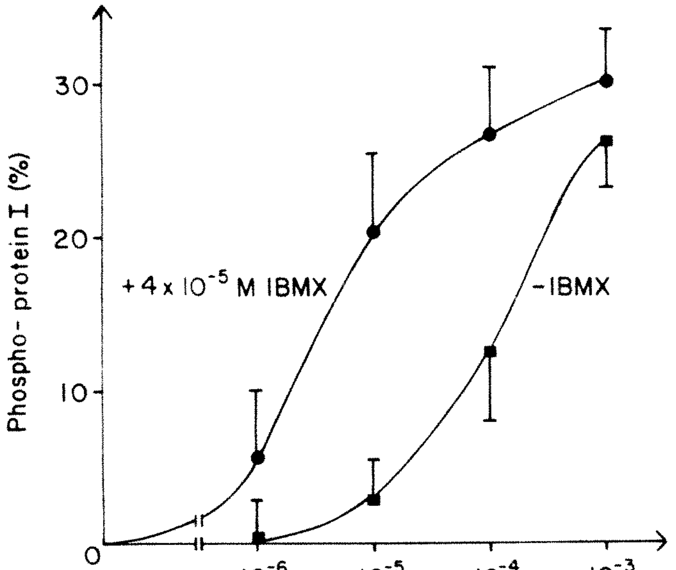


Fig. 2 Formation of phospho-Protein I from dephospho-Protein I in slices of rat facial nucleus incubated in the presence of various concentrations of serotonin. Experimental slices were incubated in the presence of the indicated concentration of serotonin, in the presence (●) or absence (■) of 4×10^{-5} M IBMX. Paired control slices were incubated in the absence of serotonin, in the presence or absence of 4×10^{-5} M IBMX, respectively. After termination of the incubation and extraction of Protein I from the slices, the amount of Protein I present in the dephospho-form in the extract was assayed by back-phosphorylation of dephospho-peptide 1, as described in Fig. 1 legend. The amount of phosphate incorporated into Protein I was determined as previously described[8], and the per cent conversion of dephospho- to phospho-Protein I due to serotonin was then calculated for each pair of incubations. Each data point represents the mean ± s.e.m. of 5–10 pairs of samples, each assayed in duplicate. The basal level of phosphate incorporation into Protein I was 7.9±0.8 pmol per mg Zn acetate-precipitable protein ($n=25$) in the absence of IBMX, and 7.4±0.7 pmol per mg Zn acetate-precipitable protein ($n=25$) in the presence of IBMX (4×10^{-5} M).

In a previous study of the facial nucleus, using electron microscope autoradiography with ^{3}H-serotonin, no specialized axo-axonic synapses involving serotonergic terminals were seen[13]. All the serotonergic terminals observed to form synapses innervated motoneurone somata or dendrites[13]. Electrophysiological studies suggest that a major effect of the serotonin input to the facial nucleus is to facilitate motoneurone excitation by a direct action on the motoneurone cell bodies[10,13,27]. The present studies on Protein I phosphorylation indicate that serotonin may also serve a paracrine function[28], diffusing to neighbouring presynaptic elements and possibly modifying neurotransmitter release. Thus, in the facial motor nucleus, serotonin may modulate function in both pre- and postsynaptic elements. Support for this interpretation is provided by studies of the crustacean neuromuscular junction, where serotonin acts both on excitatory nerve endings to facilitate transmitter release[29] and directly on muscle fibres to stimulate contracture[30].

We thank Dr G. K. Aghajanian for many helpful discussions and for provision of unpublished data. This work was supported by USPHS grants AA-04183, DA-01627, MH-17387 and NS-08440 and a grant from the McKnight Foundation.

Received 8 July; accepted 10 November 1980.

1. Ueda, T. & Greengard, P. *J. biol. Chem.* **252**, 5155–5163 (1977).
2. Ueda, T., Maeno, H. & Greengard, P. *J. biol. Chem.* **248**, 8295–8305 (1973).
3. Krueger, B. K., Forn, J. & Greengard, P. *J. biol. Chem.* **252**, 2764–2773 (1977).
4. Huttner, W. B. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5402–5406 (1979).
5. Bloom, F. E., Ueda, T., Battenberg, E. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5982–5986 (1979).
6. De Camilli, P., Ueda, T., Bloom, F. E., Battenberg, E. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5977–5981 (1979).

7. Ueda, T. *et al. J. Cell Biol.* **83,** 308–319 (1979).
8. Forn, J. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5195–5199 (1978).
9. Courville, J. *Brain Res.* **1,** 338–354 (1966).
10. McCall, R. B. & Aghajanian, G. K. *Brain Res.* **169,** 11–27 (1979).
11. Dahlström, A. & Fuxe, K. *Acta physiol. scand.* **64,** 1–36 (1965).
12. Palkovits, M., Brownstein, M. & Saavedra, J. M. *Brain Res.* **80,** 237–249 (1974).
13. Aghajanian, G. K. & McCall, R. B. *Neuroscience* (in the press).
14. Brown, B. L., Dolphin, A. C. & Greengard, P. (in preparation).
15. Greengard, P. *Harvey Lect.* (Academic, New York, 1981).
16. Reddington, M., Rodnight, R. & Williams, M. *Biochem J.* **132,** 475–482 (1973).
17. De Blas, A. L., Wang, Y. J., Sorensen, R. & Mahler, H. R. *J. Neurochem.* **33,** 647–659 (1979).
18. Ng, M. & Matus, A. *Neuroscience* **4,** 169–180 (1979).
19. Zwiers, H., Tonnaer, J., Wiegant, V. M., Schotman, P. & Gispen, W. H. *J. Neurochem.* **33,** 247–256 (1979).
20. Lohmann, S. M., Walter, U. & Greengard, P. *J. biol. Chem.* (in the press).
21. Vargaftig, B. B., Coignet, J. L., De Vos, C. J., Grijsen, H. & Bonta, I. L. *Eur. J. Pharmac.* **16,** 336–346 (1971).
22. Nelson, D. L., Herbert, A., Bourgoin, S., Glowinski, J. & Hamon, M. *Molec. Pharmac.* **14,** 983–995 (1978).
23. Enjalbert, A., Hamon, M., Bourgoin, S. & Bockaert, J. *Molec. Pharmac.* **14,** 11–23 (1978).
24. Swanson, L. W. & Hartman, B. K. *J. comp. Neurol.* **163,** 467–506 (1975).
25. Dolphin, A. C. & Greengard, P. *J. Neurosci.* (in the press).
26. Whitman, M., Dolphin, A. C., Huttner, W. B. & Greengard, P. (in preparation).
27. McCall, R. B. & Aghajanian, G. K. *Neuroscience* **4,** 1501–1510 (1979).
28. Reader, T. A., Ferron, A., Descarries, L. & Jasper, H. H. *Brain Res.* **160,** 217–229 (1979).
29. Dudel, J. *Naunyn-Schmiedebergs Archs exp. Path. Pharmak.* **249,** 515–528 (1965).
30. Kravitz, E. A., Battelle, B.-A., Evans, P. D., Talemo, B. R. & Wallace, B. G. *Neurosci. Symp.* **1,** 67–81 (1975).
31. Bradford, M. M. *Analyt. Biochem.* **72,** 248–254 (1976).
32. Beavo, J. A., Bechtel, P. J. & Krebs, E. G. *Meth. Enzym.* **38C,** 299–309 (1974).
33. Huttner, W. B., DeGennaro, L. J. & Greengard, P. *J. biol. Chem.* (in the press).

Defect in cyclic AMP phosphodiesterase due to the *dunce* mutation of learning in *Drosophila melanogaster*

Duncan Byers*

Division of Biology, California Institute of Technology, Pasadena, California 91125

Ronald L. Davis* & John A. Kiger Jr

Department of Genetics, University of California at Davis, Davis, California 95616

Cyclic AMP is an intracellular mediator ('second messenger') in the nervous and endocrine control of cellular function, regulating different processes in different cell types. Although evidence is incomplete, it seems that cyclic AMP enhances the calcium-mediated release of neurotransmitter in some neurones[1–3]. A simple form of memory in the mollusc *Aplysia* is probably encoded as a cyclic AMP-induced enhancement of neurotransmission at certain synapses of the central nervous system[4]. The possibility that cyclic AMP participates in learning mechanisms may be explored using genetic mutants. For this purpose the fruitfly *Drosophila* is suitable as it is genetically well characterized and can learn through olfaction[5], vision[6] or taste[7]. We show here that independent searches for mutations of olfactory learning[8] and of cyclic AMP metabolism[9], and for mutations causing female infertility[10] have each led to the same gene—the *dunce* gene. Our evidence indicates that the normal *dunce* gene may specify a cyclic AMP phosphodiesterase.

Two different mutant alleles of the *dunce* gene, *dunce*^{1} and *dunce*^{2}, were isolated in a search for mutations of olfactory learning[8]. The *dunce* mutants do not learn to avoid an odorant associated with electric shock or quinine powder, although both of these are effective reinforcers for normal flies. The mutants can sense odorants, however, as they show normal inherent olfactory preferences in other situations[8,11]. The learning performance of *dunce* in other sorts of learning test can be normal (visual learning[12]), intermediate (leg-position learning; R. Booker and W. G. Quinn, personal communication) or poor (larval olfactory learning[13]). In morphology and general vigour, *dunce*^{1} and *dunce*^{2} seem normal. Both mutations reduce the fecundity of females, *dunce*^{1} to about half the normal number of offspring and *dunce*^{2} to 2% of the normal[14].

Table 1 Learning in the *dunce* mutants

Genotype	Olfactory learning index (Λ)
+ (normal Canton-S)	0.33 ±0.01 ($n = 463$)
dunce^1	0.04 ±0.01 ($n = 88$)
dunce^2	0.05 ±0.01 ($n = 47$)
Df(1)N$^{71h24\text{-}5}$/*Df(1)dm*^{75e19}	0.03 ±0.03 ($n = 6$)
dunce^1/*dunce*^{2}	0.07 ±0.02 ($n = 6$)
dunce^1/*Df(1)N*$^{71h24\text{-}5}$	0.05 ±0.01 ($n = 4$)
dunce^2/*Df(1)N*$^{71h24\text{-}5}$	0.02 ±0.02 ($n = 6$)
dunce^2/*Df(1)dm*^{75e19}	0.02 ±0.02 ($n = 5$)
+/*dunce*^{1}	0.26 ±0.02 ($n = 9$)
+/*dunce*^{2}	0.25 ±0.05 ($n = 9$)
+/*Df(1)N*$^{71h24\text{-}5}$	0.18 ±0.03 ($n = 5$)
+/*Df(1)dm*^{75e19}	0.26 ±0.03 ($n = 5$)

Scores are mean ± s.e.m. of n independent determinations of the learning index, Λ. The mutants *dunce*^{1} and *dunce*^{2} were previously referred to as *dunce*^{DB38} and *dunce*^{DB276} respectively[11–13]. The learning tests of normal (+), *dunce*^{1} and *dunce*^{2} were made with males alone, females alone and mixed groups; there were no significant sexual differences in learning. The other genetic combinations were tested as females heterozygous for two different X chromosomes, listed one before and one after a solidus. Before learning was tested, the two deficiency strains were backcrossed for two generations with strains derived from Canton-S to reduce possible genetic variation due to chromosomes other than the X. For learning measurements, flies were trained and tested in groups as previously described[5,8], with 3-octanol and 4-methylcyclohexanol as odorants and electric shock as punishment (see text). The learning index, Λ, is defined as the fraction of flies avoiding the odorant previously associated with shock minus the fraction avoiding the control odorant[5]. A positive score indicates learning (theoretical maximum = 1.0) and zero indicates no learning.

Other mutations of the *dunce* gene and some genetic deletions of *dunce* were identified through study of the genetics of cyclic AMP metabolism in *Drosophila*. As in other animals, cyclic AMP is synthesized from ATP by adenyl cyclase and hydrolysed to 5′-AMP by cyclic AMP phosphodiesterase. Animals typically have several forms (isozymes) of cyclic AMP phosphodiesterase[15–17]. In *Drosophila*, two forms, here designated I and II, have been described and partially purified; these differ in molecular weight, substrate specificity and other properties[18,19]. A search for genes with major control over the cyclic AMP phosphodiesterase activity revealed that the 'genetic dosage' of region 3D3–3D4 of the X chromosome affects the activity of the smaller form (form II) in a manner suggesting that this region may include the structural gene of form II[9,18]. Using two genetic deletions, *Df(1)N*$^{71h24\text{-}5}$ and *Df(1)dm*^{75e19} (see Fig. 1), one can generate (female) flies not carrying this region on either X chromosome[20]. These females are morphologically

Table 2 Cyclic AMP phosphodiesterase activity in unfractionated homogenates of the *dunce* mutants

Genotype	Cyclic AMP phosphodiesterase activity (pmol cyclic AMP hydrolysed per min per mg live weight) Male	Female
Normal (Canton-S)	226±7 ($n = 10$)	174±8 ($n = 11$)
dunce^1	135±13 ($n = 6$)	113±10 ($n = 6$)
dunce^2	94±13 ($n = 6$)	101±5 ($n = 6$)

Activities are mean ± s.e.m. of n assays using the procedure of Kiger and Golanty[9]. For each assay two flies were homogenized at 0 °C in 0.4–1.0 ml of buffer. Duplicate 0.05-ml portions of each homogenate were mixed with 0.05 ml of tritium-labelled cyclic AMP in water, yielding reaction mixtures containing 10^{-4} M cyclic AMP, 10 mM $MgCl_2$, 2 mM 2-mercaptoethanol and 40 mM Tris-HCl, *p*H 8.0. The mixtures were incubated at 30 °C for 20 min and then heated at 90 °C to stop the reactions. Substrate and products were separated by chromatography on paper, eluted from the paper and measured in a scintillation counter.

---

* Present addresses: European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69 Heidelberg, FRG (D.B.); Department of Chemistry, California Institute of Technology, Pasadena, California 91125 (R.L.D.).

0028-0836/81/020079—03$01.00

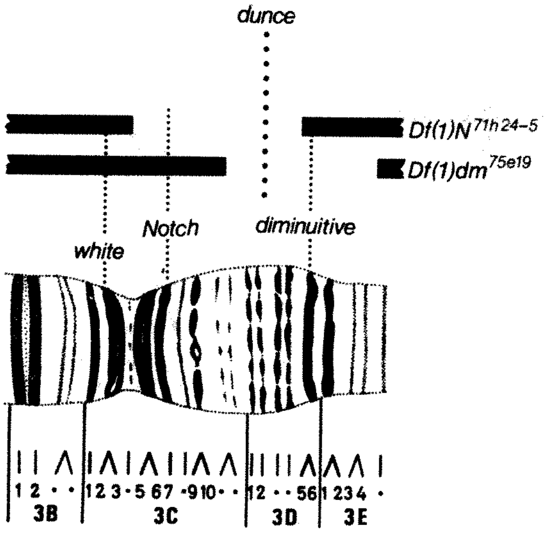


Fig. 1 Cytogenetic location of the *dunce* gene and two deletions of the *dunce* gene. At the top are illustrated the breakpoints found by Lefevre of *Deficiency (first chromosome) Notch*$^{71h24\text{-}5}$ *(Df(1)N*$^{71h24\text{-}5}$*)*, a deletion of chromosomal bands 3C4 to 3D4 inclusive, and *Df(1) diminuitive*^{75e19} *(Df(1)dm*^{75e19}*)*, a deletion of bands 3C12 to 3E4 inclusive[27]. The horizontal solid bars indicate regions of intact chromosome and the gaps indicate deleted regions. In alignment below is part (3%) of Bridges' drawing[28] of the *Drosophila* X chromosome of the larval salivary cells. Locations of the genes *white*, *Notch* and *diminuitive* are also shown.

normal and appear to be vigorous and healthy, even though they lack five adjacent chromosomal bands, including the *dunce* locus. They are sterile, and they lack the form II of cyclic AMP phosphodiesterase activity.

Still two other mutant alleles of the *dunce* gene have been found among 225 female-sterile mutants isolated by Mohler[10] on the *Drosophila* X chromosome. We show no data for them here but introduce them briefly. Both are located within chromosomal region 3C12–3D4 and are deficient in the form II of cyclic AMP phosphodiesterase; this was the first evidence that female infertility and defective cyclic AMP metabolism are associated with one gene in this region rather than two neighbouring genes[21]. In addition, the two mutants of Mohler are unsuccessful in olfactory learning and fail to complement *dunce*^{2} in learning or fertility[14]. Therefore they are alleles of the dunce gene, and we, with Mohler, rename them *dunce* ^{M11} and *dunce* ^{M14}.

Measurements of olfactory learning ability were made using the olfactory learning paradigm of Quinn, Harris and Benzer[5]. For training, groups of approximately 50 flies were attracted with a light to run alternately on to two printed circuit grids spread with different odorants, with one of the grids electrified. To test later whether learning had occurred, similar grids were used without shock, and the number of flies avoiding each odorant grid recorded. Normal and *dunce* flies tend to avoid the charged grid during training. Afterwards, however, normal flies selectively avoid the odorant previously paired with shock, whereas *dunce* flies do not[8]. Normal flies can learn to avoid the odorant on either of the two training grids, depending only on which grid is electrified: therefore the selective avoidance is not due to pseudoconditioning or inherent odour bias, but is an example of associative learning[5]. The learning index, Λ, defined in Table 1 legend, is a measure of the selective avoidance[5].

For normal flies of the Canton-S strain, Λ is 0.33 (Table 1). The mutants *dunce*^{1}, *dunce*^{2} and *Df(1)N*$^{71h24\text{-}5}$*/Df(1)dm*^{75e19} learn poorly ($\Lambda = 0.04$, 0.05 and 0.03, respectively). Heterozygous *dunce*/ + flies do not seem to learn as well as +/+ normal flies, suggesting that one normal gene may not be sufficient for normal learning. Flies heterozygous for pairs of different *dunce* deletions or mutations learn poorly (Table 1), showing that the different mutations do not complement each other in the combinations tested. Therefore, the *dunce* gene is located within the region of overlap of the two deletions.

The cyclic AMP phosphodiesterase activity of *dunce* was measured in two ways. First, adult flies were homogenized, radio-labelled cyclic AMP was added and the rate of hydrolysis of cyclic AMP measured. The rate of hydrolysis of cyclic AMP is reduced in *dunce*^{1} and *dunce*^{2} to approximately half of normal (Table 2). Second, groups of normal or *dunce* flies were homogenized, then centrifuged briefly to remove particulates, the supernatants were fractionated on sucrose density gradients and the distribution of cyclic AMP phosphodiesterase activity on the gradients measured. Normal *Drosophila* display two prominent forms (I and II) of soluble cyclic AMP phosphodiesterase[18], as shown in Fig. 2. The mutants *dunce*^{1} and *dunce*^{2} seem to have

Table 3 Cyclic AMP in the *dunce* mutants

Genotype	Cyclic AMP in adult whole fly (pmol per mg live weight) Male	 Female
Normal (Canton-S)	1.5±0.1 ($n=4$)	1.6±0.1 ($n=6$)
dunce^{1}	2.7±0.1 ($n=6$)	2.3±0.2 ($n=6$)
dunce^{2}	2.8±0.1 ($n=6$)	2.6±0.1 ($n=5$)

Cyclic AMP was measured by radioimmunoassay (NEN) in which *Drosophila* cyclic AMP competed with a standard amount of ^{125}I-labelled cyclic AMP for antibody binding sites. Values are mean ±s.e.m. of n extracts assayed. For each extract, four to six flies were homogenized together in 0.1 M HCl at room temperature. The extracts were heated for 2 min in a boiling water bath, cleared by centrifugation, neutralized to pH 6.2, cleared again by centrifugation, diluted and assayed as described previously[22]. The recovery of cyclic AMP was estimated with a small quantity of tritium-labelled cyclic AMP added to the initial homogenates and averaged 96%. If the cyclic AMP of normal flies (1.5 pmol per mg) were distributed uniformly through the fly, its concentration would be ~1.5 μM.

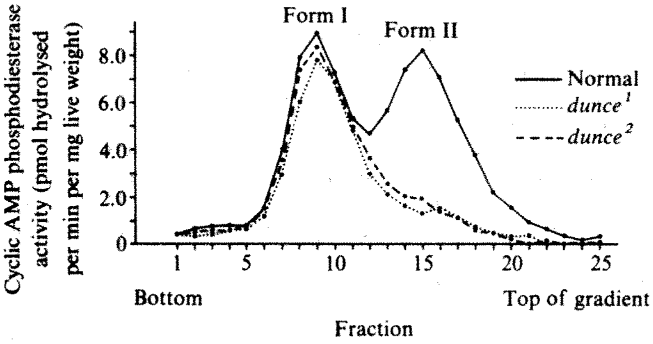


Fig. 2 Sucrose density gradient sedimentation of the soluble cyclic AMP phosphodiesterase activity of normal (Canton-S) and *dunce Drosophila*. Enzyme activity is expressed in pmol of cyclic AMP hydrolysed per min per mg live weight. Groups of 40 adult females were homogenized at 0 °C in 1.0 ml of buffer (10 mM $MgCl_2$, 1 mM $CaCl_2$, 2 mM 2-mercaptoethanol, and 40 mM Tris-HCl, pH 8.0). The same buffer was used for all subsequent steps. The homogenates were centrifuged at 105,000g for 60 min. 0.2-ml portions of the supernatants were layered on 5–20% sucrose density gradients (4.6 ml) and centrifuged at 165,000g (r_{av}) for 18 h. The gradients were divided into 0.18-ml fractions and duplicate 0.08-ml portions of each fraction were mixed with 0.02 ml of 5×10^{-4} M tritium-labelled cyclic AMP (NEN), incubated at 30 °C for 30 min, and then heated at 90 °C to stop the reactions. The product, 5'-AMP, was converted to adenosine with unpurified *Crotalus atrox* venom (Sigma) and measured with the resin assay of Thompson and Appleman[15]. Blank controls were treated identically except that buffer was added instead of homogenate; the radioactivity of the blanks (4,100 c.p.m.) was subtracted from each fraction. Before correction, the most active fractions contained about 15,000 c.p.m., representing hydrolysis of 12% of the radioactive cyclic AMP. The recovery of applied cyclic AMP phosphodiesterase activity from each gradient was ~50%.

normal amounts of form I and their activity at the position of form II is small, though not zero (Fig. 2). On similar gradients of $dunce^{M11}$, $dunce^{M14}$ and $Df(1)N^{71h24-5}/Df(1)dm^{75e19}$ (not shown), form I is approximately normal and form II is undetectable (less than 5% of normal).

If the activity of a cyclic AMP phosphodiesterase were reduced, then the concentration of cyclic AMP in the fly might be expected to rise. To test this, the amount of cyclic AMP in normal and *dunce* flies was measured using a radioimmunoassay. Normal adults yielded 1.5 pmol of cyclic AMP per mg live weight (Table 3 and ref. 22). The mutants $dunce^1$ and $dunce^2$ yielded about 160% of the normal quantity of cyclic AMP (Table 3). Cyclic AMP in similar extracts of $dunce^{M11}$, $dunce^{M14}$ and $Df(1)N^{71h24-5}/Df(1)dm^{75e19}$ is 300–600% of the normal[21,22].

We conclude that the normal *dunce* gene product has roles in behaviour, reproduction and cyclic AMP metabolism. The results show that the *dunce* gene and one isozyme of cyclic AMP phosphodiesterase are closely related in some way, but further experiments are required to test whether *dunce* is the structural gene of this enzyme. If elimination of cyclic AMP phosphodiesterase is in fact the cause of poor learning in *dunce*, then pharmacological inhibition of cyclic AMP phosphodiesterase activity should interfere with learning in normal flies. This expectation is confirmed by an experiment of Dudai and Jan[23], recently reconfirmed by Dudai[24]. They found that the olfactory learning performance of *dunce* mutants is mimicked by normal flies fed caffeine, an inhibitor of cyclic nucleotide phosphodiesterases[25,26]. One question still unanswered is whether the normal product of the *dunce* gene forms an integral part of a molecular apparatus of learning.

We thank J. D. Mohler and G. Lefevre Jr for mutant strains. D.B. thanks Seymour Benzer for support and laboratory facilities, and Lawrence Kauvar, Sandra Shotwell and Shinobu Fujita for biochemical advice. Financial support was from the NSF (to S. Benzer) and NIH (to J.A.K.); D.B. was a fellow of the Gordon Ross Medical Foundation and R.L.D. was a NIH Pre-Doctoral Trainee.

Received 22 July; accepted 25 September 1980.

1. Shimihara, T. & Tauc, L. *Brain Res.* **127,** 168-172 (1977).
2. O'Shea, M. & Evans, P. D. *J. exp. Biol.* **79,** 169–190 (1979).
3. Goldberg, A. L. & Singer, J. J. *Proc. natn. Acad. Sci. U.S.A.* **64,** 134–141 (1969).
4. Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3512–3516 (1978).
5. Quinn, W. G., Harris, W. A. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **71,** 708–712 (1974).
6. Menne, D. & Spatz, H. C. *J. comp. Physiol.* **114,** 301–312 (1977).
7. Medioni, J. & Vaysse, G. *C. r. séanc. Soc. Biol.* **169,** 1386–1391 (1975).
8. Dudai, Y., Jan, Y. -N., Byers, D., Quinn, W. G. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 1684–1688 (1976).
9. Kiger, J. A. Jr & Golanty, E. *Genetics* **85,** 609–622 (1977).
10. Mohler, J. D. *Genetics* **74,** s184 (1973); **85,** 259–272 (1977).
11. Dudai, Y. *J. comp. Physiol.* **130,** 271–275 (1979).
12. Dudai, Y. & Bicker, G. *Naturwissenschaften* **65,** 494–495 (1978).
13. Aceves-Piña, E. O. & Quinn, W. G. *Science* **206,** 93–96 (1979).
14. Byers, D. thesis, California Inst. Technol. (1979).
15. Thompson, W. J. & Appleman, M. M. *Biochemistry* **10,** 311–316 (1971).
16. Van Imwegen, R. G., Pledger, W. J., Strada, S. J. & Thompson, W. J. *Archs Biochem. Biophys.* **175,** 700–709 (1976).
17. Weiss, B. *Adv. Cyclic Nucleotide Res.* **5,** 195–211 (1975).
18. Kiger, J. A. Jr & Golanty, E. *Genetics* 91, 521–535 (1979).
19. Davis, R. L. & Kiger, J. A. Jr *Archs Biochem. Biophys.* **203,** 412–421 (1980).
20. Kiger, J. A. Jr *Genetics* **85,** 623–628 (1977).
21. Davis, R. L. thesis, Univ. California (1979).
22. Davis, R. L. & Kiger, J. A. Jr *Biochem. biophys. Res. Commun.* **81,** 1180–1186 (1978).
23. Dudai, Y. & Jan, Y. -N. *Biol. a. Rep.*, California Inst. Technol., 72 (1975).
24. Dudai, Y. *Abstr. a. Mtg Isr. biochem. Soc.*, 100 (1979).
25. Sutherland, E. W. & Rall, T. W. *J. biol. Chem.* **232,** 1077–1091 (1958).
26. Goodman, L. S. & Gilman, A. (eds) in *Pharmacological Basis of Therapeutics* 5th edn, 367–378 (Macmillan, New York, 1975).
27. Craymer, L. & Roy, E. *Drosoph. Inf. Serv.* **55,** 200–204 (1980).
28. Bridges, C. B. *J. Hered.* **29,** 11–13 (1938).

Altered substrate specificity of herpes simplex virus thymidine kinase confers acyclovir-resistance

G. Darby & H. J. Field

Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 2QQ, UK

S. A. Salisbury

Department of Organic and Inorganic Chemistry, University of Cambridge, Cambridge, UK

Acyclovir (9-[2-hydroxyethoxymethyl]guanine or ACV) is a nucleoside analogue with considerable potential for the treatment of herpes simplex virus (HSV) infections in man[1,2]. Two virus-coded enzymes are important in the mechanism of action of this drug: thymidine kinase (TK) which initiates its activation by converting it to the monophosphate[3] and DNA polymerase whose action is inhibited by ACV triphosphate[4]. Changes in either gene may confer resistance[5–7], but all reported mutations in the TK gene have resulted in failure of the resistant virus to induce appreciable levels of the enzyme. Such TK$^-$ mutants arise readily in tissue culture systems[8,9] where the enzyme is non-essential for virus replication[10], but in animals they show considerably reduced pathogenicity and neurovirulence[7,11–14]. We now describe the isolation of a resistant mutant which induces a TK of altered substrate specificity and we show that this virus retains pathogenicity for mice with only a slight attenuation of neurovirulence.

To obtain conditions *in vitro* which would be selective against the readily occurring TK$^-$ variants, serum-starved BHK cells were used. It is known that TK$^-$ virus is at a disadvantage in resting cells[11,15] and it was therefore hoped that selecting for resistance in these cells would increase the probability of isolating TK$^+$ resistant virus. In fact both TK$^+$ and TK$^-$ resistant clones were obtained using this method.

Table 1 Resistance of TK$^+$ and TK$^-$ variants of SC16

Virus	TK (% SC16)	Resistance to ACV (ED_{50}, μg ml^{-1}) BHK cells	$D2_1$ cells*	Resistance to PAA (ED_{50}, μg ml^{-1}) BHK cells
SC16	100	0.02	0.01	11
SC16 S1	30–40	2.7	0.01	12
SC16 S6	<0.1	3.0	ND	ND

The parental strain of HSV1 (SC16) was originally isolated from an oral lesion. It has been well characterized in mice[21] and has previously been used to derive ACV-resistant variants[7]. 2×10^6 BHK cells were seeded into 25-cm^2 plastic tissue-culture flasks and incubated overnight in Glasgow-modified Eagle's medium supplemented with 10% tryptose phosphate broth and 10% calf serum. After 24 h the medium was removed, the cells washed and the flasks replenished with Eagle's medium without tryptose phosphate broth and with calf serum reduced to 0.5%. The cells were then maintained for 5 d before being inoculated with SC16 at a multiplicity of 0.1 plaque-forming unit (PFU) per cell in the presence of 0.1 μg ml^{-1} ACV. After 3 d of incubation, the cells were collected, subjected to ultrasonic disruption, then the released virus was titrated and used to inoculate further bottles of serum-starved cells. The second and third passage were in 1.0 and 10 μg ml^{-1} ACV respectively. Clones were isolated from the virus yield from the third passage. TK activity was measured on extracts of BUdR-resistant BHK (BU-BHK) cells obtained 18 h after inoculation with virus at 5 PFU per cell. The cells were resuspended at 3×10^7 per ml in 0.01M Tris and disrupted by ultrasonic vibration. Reactions were performed at 37 °C in 0.125-ml aliquots containing 0.02M phosphate buffer pH 6, 5mM $MgCl_2$, 5mM ATP and 40 μM 14C-thymidine (1.6 μCi ml^{-1}). Reactions were carried out using the method described by Klemperer *et al.*[22]. The results are expressed as % phosphorylation compared with the SC16 extract. Plaque reduction assays for both ACV and PAA were performed by inoculating monolayers of BHK cells with 200 PFU virus. After adsorption for 1 h, overlay medium was added containing increasing concentrations of the test drug. Graphs were plotted of % plaque reduction against $\log_{10}$ drug concentration and the ED_{50} values were assessed directly from the graphs.

* $D2_1$ cells are LTK$^-$ cells transformed to TK$^+$ phenotype using sheared *HSV-2 DNA*[23].

0028-0836/81/020081—03$01.00

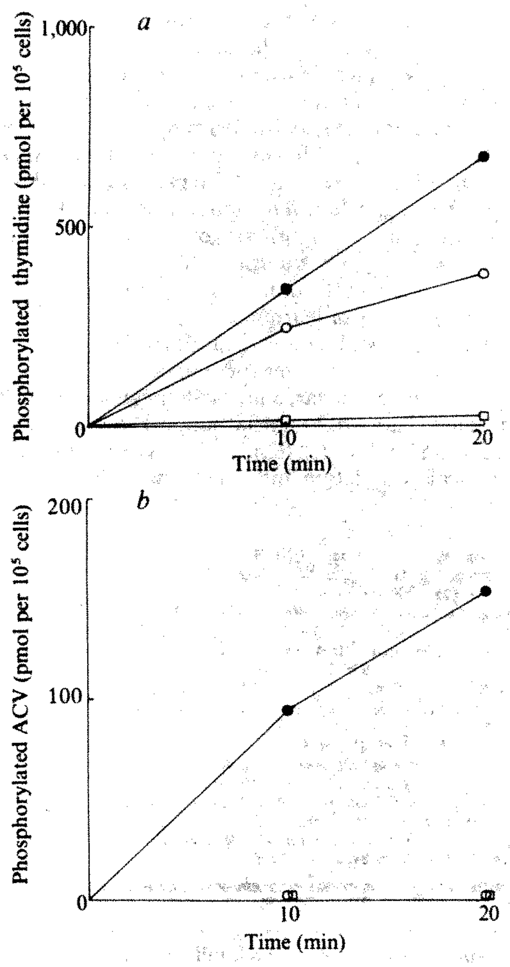


Fig. 1 The phosphorylation of thymidine and acyclovir by extracts of BU BHK cells infected with wild-type or ACV-resistant HSV. *a*, Phosphorylation of thymidine; *b*, phosphorylation of acyclovir. ●, SC16; ○, SC16 S1; □: SC16 S6. The preparation of infected cell extracts and assay for thymidine kinase are described in Table 1 legend. The assay for ACV was similar except that ^{14}C-thymidine was replaced by 0.8 mM ^{3}H-ACV (40 μCi ml^{-1}). The reactions were allowed to proceed for 0, 10 and 20 min and the 0-min values were subtracted. ^{3}H-ACV was prepared by a modification of the procedure described by Shelton and Clark[24].

HSV-1 strain SC16 was passaged in the presence of increasing concentrations of ACV in serum-starved BHK cells and the yield from the third passage was cloned in the absence of drug in actively growing BHK cells. Two of these clones (SC16 S1 and SC16 S6) were examined in detail. Both were shown to be approximately 100-fold more resistant to ACV (ED$_{50}$ 2–3 μg ml^{-1}) than the parental strain (Table 1). Thymidine kinase assays on infected cell extracts taken 18 h after high multiplicity infection showed that SC16 S1 induced TK whereas no activity was detected in the case of S6. However, the level of TK induced by SC16 S1 was rather less (30–40%) than that induced by the parental virus. One possible explanation for this difference would be a mutation in the TK gene.

Three lines of evidence were subsequently obtained which also suggested that the resistance of SC16 S1 resides in an altered TK and not at the level of DNA polymerase. First, SC16 S1 was sensitive to ACV when tested in D2$_1$ cells which carry and express the HSV TK gene (Table 1). These cells 'activate' the drug and only viruses which have acquired resistance at loci other than TK are resistant in them[8,16]. Secondly, viruses which are resistant at the DNA polymerase locus are also frequently resistant to phosphonoacetic acid (PAA)[5,6,8]. However, SC16 S1 was as sensitive as the parent to this drug (Table 1).

The third line of evidence concerns the direct measurement of the ability of the induced TK to phosphorylate ACV *in vitro*. The ability of extracts from cells infected with SC16 or ACV-resistant strains to phosphorylate thymidine and ACV were compared. The results of this experiment are shown in Fig. 1. Clearly the parent virus is able to phosphorylate both thymidine

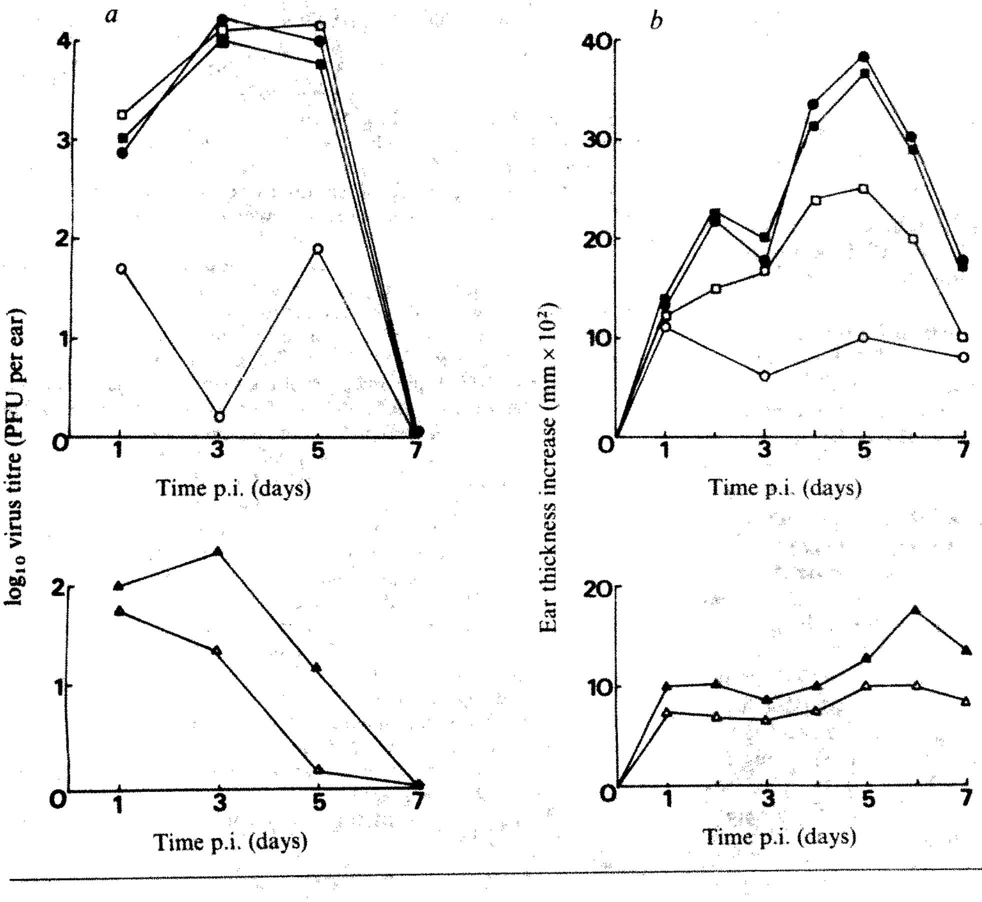


Fig. 2 Pathogenicity for mice of wild-type and TK^{+}- and TK^{-}-resistant strains of HSV, with and without ACV-treatment. *a*, Virus growth; *b*, cell-mediated inflammation. SC16: ●, untreated; ○, ACV-treated. SC16 S1: ■, untreated; □, ACV-treated. SC16 S6: ▲, untreated; △, ACV-treated. 10^{4} PFU test virus were inoculated into the left ear pinna of anaesthetized, 4-week-old female BALB/c mice. Virus growth in the ear was measured by homogenizing the pinnae from groups of three mice, and virus titres were determined by plaque titration in BHK cells. Ear thickness was measured using an engineer's micrometer—the thickness increase has been found to be a reliable indicator of cell-mediated inflammation[18]. The drug was suspended in water and given in twice daily intraperitoneal injections totalling 50 mg per kg per day. The effects of ACV on sensitive and resistant strains of HSV in this model have been described previously[7,17]. p.i., Post-infection.

Table 2 Virulence of TK^+ and TK^- variants of SC16 in BALB/c mice

Virus	TK induction	Ear inoculation latency in cervical ganglia*	i.c. Inoculation PFU/LD_{50}†
SC16	+	7/10 (0/10)	6
SC16 S1	+	5/12	60
SC16 S6	–	0/12	$>10^6$

* These mice are the survivors from the experiment illustrated in Fig. 2. The second, third and fourth cervical ganglia were explanted into Eagle's medium and incubated at 37 °C for 6 d. The tissue was then homogenized and tested for the presence of infectious virus by plaque titration in BHK cells as described previously[11]. Ratios are the number of mice from which latent virus was reactivated over the number of animals tested, the number in parentheses being the proportion positive following twice daily treatment with 50 mg per kg per day ACV injected intraperitoneally from the time of virus inoculation until 10 d thereafter.

† Dilutions of virus were inoculated into the left cerebral hemisphere in 0.02 ml Eagle's medium. The ED_{50} value was obtained using the Spearman–Karber method from the numbers of mice dying between days 2 and 14 after inoculation.

and ACV, though the actual amount of ACV phosphorylated is very much less. In contrast, the mutant-infected cell extracts failed to phosphorylate ACV although the S1 extract contained significant TK activity. Another isolate of HSV [C1(101)] which induces similar levels of TK to SC16 S1 was also tested; in this case ACV-phosphorylating activity was readily measured[20]. All the evidence is therefore consistent with the resistance of SC16 S1 residing in its failure to activate the drug in infected cells and this in turn is presumably due to a mutation in the TK gene which generates an enzyme with an altered substrate specificity.

As the resistant virus SC16 S1 was TK^+, it was obviously of interest to investigate its virulence in an animal model system. We have previously used the mouse ear model to investigate the pathogenicity of mutant viruses[7,11]. It can be seen from the data (Fig. 2) that SC16 S1 grows as well in the mouse ear as does the parent strain. In contrast, the growth in the ear of the TK^- mutant SC16 S6 was greatly impaired. Furthermore, a normally effective treatment routine[7,17] had little inhibitory effect on this virus because there was no marked reduction in virus titres in ear tissue or inflammation as determined by measurement of ear thickness[18]. The cervical dorsal root ganglia were later explanted from surviving mice to test for the presence of a latent infection (Table 2). Latent infections were detected in ganglion explants from both SC16 and SC16 S1 but not from S6. This was to be expected because of the reduced replication of the latter virus in the skin during the acute phase of the infection.

A further measure of the neurovirulence of the viruses was obtained by measuring the ED_{50} dose following intracerebral inoculation[11,19]. SC16 S1 seemed to be slightly attenuated relative to its parent showing a 10-fold increase in the ratio of plaque-forming units per LD_{50} (Table 2). However, the TK^- strain, SC16 S6, was highly attenuated showing >1,000-fold increase. The pathogenicity data reported here are very similar to those obtained with another TK^+ ACV-resistant strain of HSV [C1(101)P_2C_5][7], though recent evidence regarding this virus strongly suggests that, unlike SC16 S1, its resistance resides at the level of DNA polymerase and not TK[20].

The results described here suggest that HSV may acquire resistance to ACV by a mutation in the TK gene which alters the substrate specificity of the induced enzyme. Both this kind of mutant and the TK^+ DNA-polymerase mutants which have already been reported would seem likely to arise *in vivo* as there is little reduction in their growth potential. It remains to be seen, however, whether such viruses will be isolated from or constitute a problem in patients undergoing treatment with nucleoside analogues such as ACV.

This work was supported by a MRC programme grant. We thank Mrs Elizabeth Lay for her technical assistance and Mrs M. Wright for preparing the manuscript. Acyclovir was a gift from Dr G. B. Elion.

Received 14 August, accepted 30 October 1980.

1. Elion, G. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **74,** 5716–5720 (1977).
2. Schaeffer, H. J. *et al. Nature* **272,** 583–585 (1978).
3. Fyfe, J. A., Keller, P. M., Furman, P. A., Miller, R. L. & Elion, G. B. *J. biol. Chem.* **253,** 8721–8727 (1978).
4. Furman, P. A. *et al. J. Virol.* **32,** 72–77 (1979).
5. Coen, D. M. & Schaffer, P. A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2265–2269 (1980).
6. Schnipper, L. E. & Crumpacker, C. S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2270–2273 (1980).
7. Field, H. J. & Darby, G. *Antimicrob. Ag. Chemother.* **17,** 209–216 (1980).
8. Field, H. J. Darby, G. & Wildy, P. *J. gen. Virol.* **49,** 115–124 (1980).
9. Smith, K. O., Kennell, W. L., Poirier, R. H. & Lynd, F. T. *Antimicrob. Ag. Chemother.* **17,** 144–150 (1980).
10. Dubbs, D. R. & Kit, S. *Virology* **22,** 439–502 (1964).
11. Field, H. J. & Wildy, P. *J. Hyg., Camb.* **81,** 267–277 (1978).
12. Tenser, R. B. & Dunstan, M. E. *Virology* **99,** 417–422 (1979).
13. Tenser, R. B., Miller, R. L. & Rapp, F. *Science* **205,** 915–917 (1979).
14. Klein, R. J., De Stephano, E., Brady, E. & Friedman-Kien, A. E. *Arch. Virol.* (in the press).
15. Jamieson, A. T., Gentry, G. A. & Subak-Sharpe, J. H. *J. gen. Virol.* **24,** 465–480 (1974).
16. Darby, G., Larder, B. A., Bastow, K. F. & Field, H. J. *J. gen. Virol.* **48,** 451–454 (1980).
17. Field, H. J., Bell, S. E., Elion, H. B., Nash, A. A. & Wildy, P. *Anitmicrob. Ag. Chemother.* **15,** 554–561 (1979).
18. Nash, A. A., Field, H. J. & Quartey-Papafio, R. *J. gen. Virol.* **48,** 351–364 (1980).
19. Marcialis, M. A. *et al. Experimenta* **31,** 502–503 (1975).
20. Field, H. J., McMillan, & Darby, G. *J. inf. Dis.* (in the press).
21. Hill, T. J., Field, H. J. & Blyth, W. A. *J. gen. Virol.* **28,** 341–353 (1975).
22. Klemperer, H. G., Haynes, G. R., Shedden, W. I. H. & Watson, D. H. *Virology* **31,** 502–503 (1967).
23. Minson, A. C., Wildy, P., Buchan, A. & Darby, G. *Cell* **13,** 581–587 (1978).
24. Shelton, K. R. & Clark, J. M. Jr, *Biochemistry* **6,** 2735–2739 (1967).

Terminal direct repeats in a retrovirus-like repeated mouse gene family

Eli Keshet & Yosef Shaul

Department of Virology, Weizmann Institute of Science, Rehovot, Israel

The mouse genome contains multiple copies of a dispersed gene family known individually as VL30 genes[1] which are thought to be associated with retroviruses. The copies consist of closely related 5.2-kilobase DNA sequences flanked by unrelated cellular DNA, and although no sequence homology has been found between the genes and the exogenous or endogenous retroviruses so far tested, the 30S RNA transcripts expressed by the genes are efficiently packaged into virions, recovered from infected cells[2–4] and transmitted to other cells by pseudo-type infection[5]. Stimulated by recent reports[6–10] of the similarity between retroviruses and transposons (from which retroviruses may have evolved[6]), and in particular by the recognition that both types of genetic elements are characterized by a large terminal repeat (LTR), we set out to determine whether the VL30 genes are also distinguished by this property. Using cloned DNAs from a mouse gene library and heteroduplex analysis, we have now found that the VL30 genes do indeed carry terminal direct repeats 400 base pairs long.

Individual copies of these 'retrovirus-like' genes (referred to as VL30 genes) were cloned from a gene library of the inbred mouse strain BALB/c in bacteriophage λ (ref. 1). The construction of the gene library from mouse genomic fragments obtained by a partial *Eco*RI digestion enabled us to select clones which, despite having *Eco*RI cleavages sites within VL30 DNA, encompass the entire 5.2 kilobases of the element and also contain some flanking DNA sequences. Such a clone with an *Eco*RI site, roughly in the middle of VL30 DNA, was used in the experiment designed to visualize LTRs (Fig. 1). As illustrated in Fig. 1, the presumptive LTRs are readily dissociated from each other by *Eco*RI digestion and subcloning of each fragment into the *Eco*RI site of plasmid pBR322. The subcloning procedure results in recombinant plasmids harbouring the eukaryotic insert in each of the two possible orientations. The orientation of insertion of several recombinant plasmids was determined by restriction enzyme digestions and blot hybridization analyses of the vector–insert junction fragments (data not shown). Two recombinant plasmids, whose VL30 DNA is inserted in the orientation depicted in Fig. 1, were linearized by *Sal*I digestion and analysed by heteroduplexing. This procedure visualizes large repeats in VL30 DNA that are in a direct orientation. A

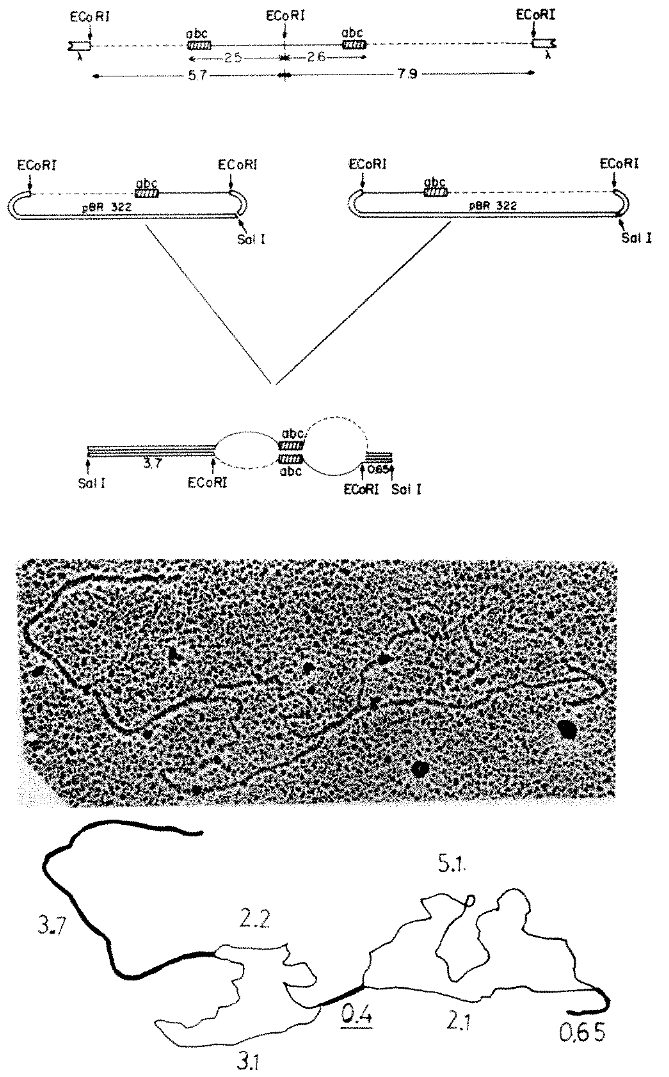


Fig. 1 Visualization of a direct repeat by heteroduplex analysis. Top: diagrammatic representation of cloned DNAs used and their expected heteroduplex. Upper drawing is of clone 3 of VL30 DNA in bacteriophage λ. Sizes (in kilobases) of the retrovirus-like DNA sequences (solid line) and their flanking DNA sequences (dashed line) were determined previously[1]. Hatched boxes indicate presumptive direct repeats in DNA. The middle drawing is of subclones of the 5.7- and 7.9-kilobase *Eco*RI fragments of clone 3 in plasmid pBR322. The lower drawing is an expected heteroduplex structure between the two *Sal*I-linearized plasmids (drawing is out of scale). Bottom: electron micrograph of a typical heteroduplex. The indicated sizes (kilobases) are average values obtained from measurements of 10 heteroduplexes. The size of the duplexed region was standardized with respect to the duplexed regions of the vector (3.7 and 0.65 kilobases). The sizes of the single-stranded regions were standardized with respect to a single-stranded ϕX174 included in the same field.

duplexed region, representing a direct repeat of ~400 base pairs, was observed (Fig. 1). A control heteroduplex, where the 5.7-kilobase *Eco*RI fragment is inserted into the plasmid in the opposite orientation (and is expected, therefore, to visualize inverted repeats) failed to reveal such a structure (data not shown). Several heteroduplexes were measured. On the basis of the known map location of the *Eco*RI site and the known sizes of the DNA sequences flanking this particular VL30 DNA[1], we could identify the direct repeat with the 400 base pairs residing at each terminus of VL30 DNA (Fig. 1).

As the inherent limitations in the accuracy of heteroduplex measurements prevent exact localization of the repeated DNA sequences within the cloned DNAs, we used another approach that exploits the availability of the poly(A)-containing 30S RNA transcripts of VL30 DNA. A hybridization probe corresponding to the 3′ end of the 30S RNA was prepared (Fig. 2). Although this probe contains sequences confined to the 3′ end—one-fifth of the RNA—it efficiently hybridized to both *Eco*RI fragments (Fig. 2), thus confirming the existence of repeated DNA sequences, and ruling out the possibility that the repeats are contained within DNA sequences adjacent to VL30 DNA. Assuming a gross co-linearity between VL30 DNA and its 30S RNA transcript, this result also places the repeated sequences at both termini of VL30 DNA.

During the propagation of recombinant phages in bacteria, some of the cloned DNAs accumulated a significant fraction of phages with a shorter eukaryotic insert. Heteroduplex analysis revealed that all these phages suffered a unique deletion of 4.7 kilobases in VL30 DNA (data not shown) and suggested to us that the deletion mutants might have arisen as a result of intramolecular recombination involving the direct terminal repeats. Such an event would lead to the loss of all VL30 DNA sequences except for one copy of the LTR. To test this possibility, a deleted form of one of the cloned phages was heteroduplexed with non-deleted phage of another VL30 clone. The rationale for this experiment is given in Fig. 3. Two types of heteroduplex are anticipated which differ with respect to the choice of the LTR copy of the non-deleted phage that anneals with the sole LTR copy retained by the deleted phage. In each case, the two heteroduplexed DNAs shared only their λ 'arms' and an additional 0.4-kilobase segment (Fig. 3). The observed heteroduplexes were consistent with structures predicted on the basis of the known size of each flanking DNA sequence and thus verified the structure suggested for the deletion. It is likely that deletions, similar to those described above, that occur during

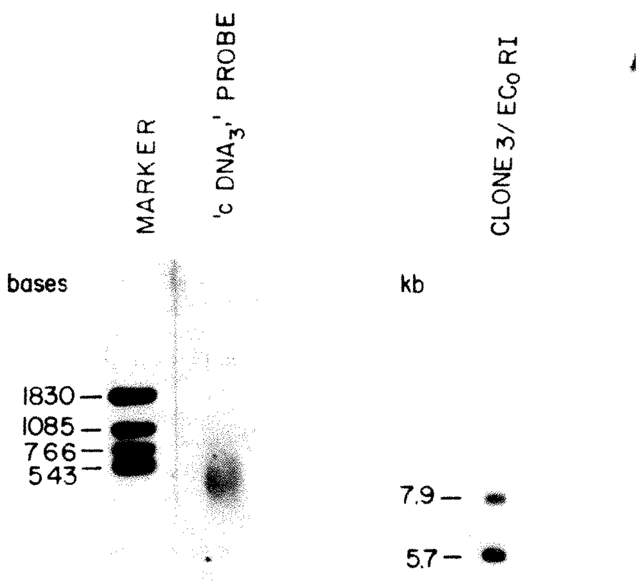
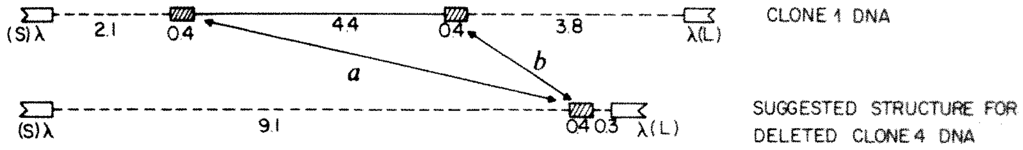


Fig. 2 Evidence for LTR in VL30 DNA with the aid of a 3′ end-specific probe. 30S RNA was extracted from purified virions of Ra-MuLV produced by JLS-V9 cells. RNA was randomly sheared and poly(A)-containing RNA was purified by two cycles of oligo(dT)-cellulose chromatography. RNA was size fractionated by centrifugation through a sucrose gradient and molecules smaller than 800 bases were pooled. ^{32}P-labelled cDNA was synthesized with oligo(dT)$_{12-18}$ as a primer. This cDNA was designated cDNA$_{3'}$. Left: size determination of the cDNA$_{3'}$ probe. A sample of cDNA$_{3'}$ was electrophoresed through a 1.2% agarose gel and autoradiographed. Size markers are ^{32}P-labelled *Hin*fI fragments of SV40 DNA. Right: 0.5 µg DNA of clone 3 DNA was digested with *Eco*RI, electrophoresed through a 0.7% agarose gel, blotted on to a nitrocellulose filter, hybridized with cDNA$_{3'}$ and autoradiographed.



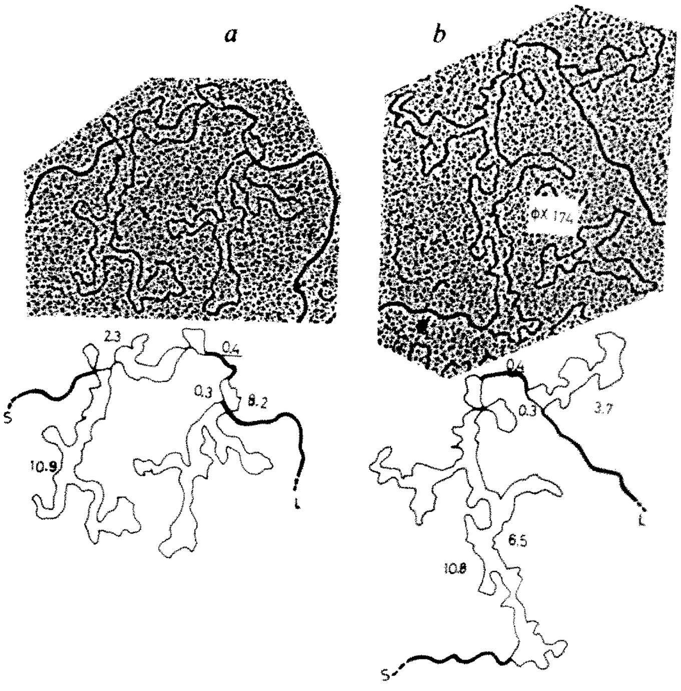


Fig. 3 Characterization of the specific deletion in cloned VL30 DNA. Top: clones 1 and 4 were previously characterized with respect to the size of VL30 DNA (solid line) and flanking DNA sequences (dotted line)[1]. Boxes indicate direct terminal repeats. S and L indicate short and long arms of the λ vector, respectively. *a* And *b* indicate the two possible forms of heteroduplexes. Bottom: electron micrographs of heteroduplexes between clone 1 DNA and deleted clone 4 DNA. A representative heteroduplex of each type is shown.

propagation of cloned VL30 DNA in bacteria, also occur in the mouse genome. Lauer *et al.* have recently described deletions in human α-globin genes that are indistinguishable from deletions that occur in bacteria with respect to the location of the breakpoints of deletion[11]. The availability of cloned LTR DNA sequences, devoid of other VL30 DNA sequences, provides a probe for such genomic deletions.

The biological significance of VL30 DNAs is unknown. They may represent degenerative endogenous viral genomes or may be cellular genetic elements of a protoviral nature. Structurally, these genetic elements strongly resemble eukaryotic transposons[12,13]. Shared features include similar size and reiteration frequencies, a dispersed genomic distribution, efficient transcription of a poly(A)-containing RNA and the presence of large direct repeats at the termini of these elements. Whether these DNA sequences actually transpose remains to be determined.

This work was carried out in the laboratory of Dr H. Aviv whom we thank for his help. We also thank Ms Y. Bernstein for technical assistance.

Received 20 August; accepted 20 October 1980.

1. Keshet, E., Shaul, Y., Kaminchik, Y. & Aviv, H. *Cell* **20,** 431–439 (1980).
2. Hawk, R. S., Troxler, D. H., Lowy, D., Duesberg, P. H. & Scolnick, E. M. *J. Virol.* **25,** 115-123 (1978).
3. Sherwin, S. A., Rapp, U. R., Benveniste, R. E., Sen, A. & Todaro, G. J. *J. Virol.* **26,** 257-264 (1978).
4. Besmer, P., Olshevsky, A., Baltimore, D., Dolberg, D. & Fan, H. *J. Virol.* **29,** 1168-1176 (1979).
5. Scolnick, E. M., Vaas, W. C., Hawk, R. S. & Duesberg, P. H. *J. Virol*, **29,** 964–972 (1979).
6. Shimotohno, K., Mizutani, S. & Temin, H. M. *Nature* **285,** 550–554 (1980).
7. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3937-3941 (1980).
8. Sutcliffe, J. G., Shinnick, T. H., Verma, I. M. & Lerner, R. A. *Proc. Natn. Acad. Sci. U.S.A.* **77,** 3302-3306 (1980).
9. Shoemaker, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 3932-3936 (1980).
10. Majors, J. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
11. Lauer, J., Shen, C.-K. J. & Maniatis, T. *Cell* **20,** 119-130 (1980).
12. Cameron, J. R., Loh, E. Y. & Davis, R. W. *Cell* **16,** 739-751 (1979).
13. Finnegan, D. S., Rubin, G. M., Young, M. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42,** 1053-1063 (1977).

A general method for site-directed mutagenesis in prokaryotes

Gary B. Ruvkun & Frederick M. Ausubel

Cellular and Developmental Biology Group, Department of Biology, Harvard University, Cambridge, Massachusetts 02138

The genetic analysis of genes from prokaryotic species for which experimental genetic systems have not yet been developed is often limited by the difficulty of producing mutations in those genes. We report here a general technique applicable to Gram-negative prokaryotes for site-directed mutagenesis of cloned DNA fragments which we have applied to the study of the symbiotic nitrogen fixation genes of *Rhizobium meliloti*. In particular, we mutagenized cloned *R. meliloti* restriction fragments in *Escherichia coli* with transposon Tn*5* and then replaced the wild-type parental DNA sequences with the mutant DNA sequences in the *R. meliloti* genome. Using this method we show that an *R. meliloti* DNA restriction fragment, cloned previously on the basis of homology to *Klebsiella pneumoniae nif* genes[1], contains gene(s) essential for symbiotic nitrogen fixation. In addition, we use this method to construct a physical genetic map of a subset of the *R. meliloti nif* genes.

The bacterial species *R. meliloti* fixes nitrogen in symbiotic association with alfalfa (*Medicago sativa* L.). Establishment of symbiosis begins with the bacterial invasion of roots and culminates in the development of the root nodule, a specialized structure within which the *Rhizobium* differentiate into nitrogen-fixing bacteroids[2]. *R. meliloti* can be grown in culture apart from its host plant, but does not normally fix nitrogen in this free-living state. Therefore, the isolation and characterization of nitrogen fixation (*nif*) mutants using standard bacterial genetic techniques would require the screening of thousands of potential bacterial mutants on individual plants. To

0028-0836/81/020085—04$01.00

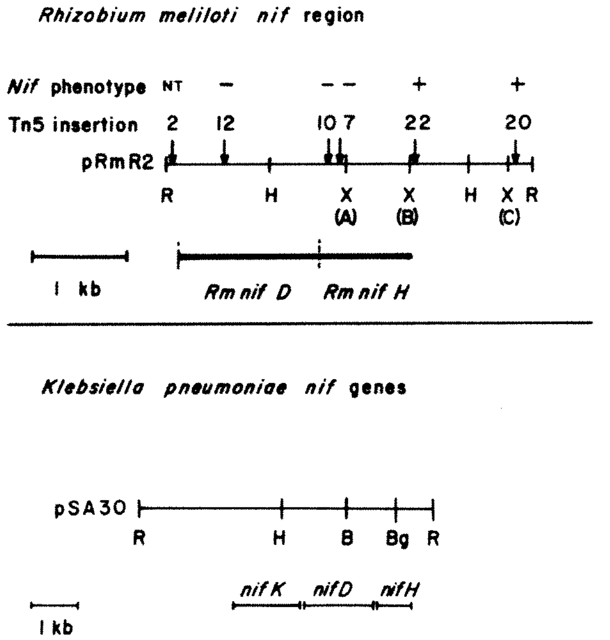

Rhizobium meliloti nif region
Nif phenotype NT − −− + +
Tn5 insertion 2 12 10 7 22 20
pRmR2
R H X X H X R
(A) (B) (C)
1 kb
Rm nif D
Rm nif H
Klebsiella pneumoniae nif genes
pSA30
R H B Bg R
nif K
nif D
nif H
1 kb

Fig. 1 Physical genetic map of *R. meliloti nif* genes *D* and *H*. Restriction maps of plasmid pRmR2 and each pRmR2::Tn*5* plasmid were constructed using a combination of digestions in 1 × universal restriction enzyne buffer (1 × UREB = 60 mM NaCl, 10 mM Tris *p*H 7.4, 6 mM $MgCl_2$, 6 mM *β*-mercaptoethanol) with restriction endonucleases *Eco*RI, *Hin*dIII, *Xho*I and *Bam*HI followed by gel electrophoresis in 1–1.5% agarose gels. These digested DNAs were transferred to nitrocellulose by the method of Southern[7] and hybridized with probes derived from the *K. pneumoniae* 6.3-kilobase *Eco*RI fragment cloned in pSA30 (ref. 22). As shown in *b*, the 1.4-kilobase *Hin*dIII–*Bam*HI fragment of pSA30 contains the C-terminal coding region of *nifD* and the 0.7-kilobase *Bgl*II–*Eco*RI fragment contains the N-terminal coding region of *nifH* (ref. 23). These fragments were eluted from preparative agarose gels as described previously[1], and labelled with [α-^{32}P]TTP (350 Ci $mmol^{-1}$, Amersham) using the nick translation method to a specific activity of 5×10^7 c.p.m. per μg of DNA. Approximately 10^6 c.p.m. were hybridized to the nitrocellulose sheet as described previously[1]. Regions of homology in the *R. meliloti nif* region to *K. pneumoniae nifD* and *nifH* genes (shown in heavy lines) were localized by a combination of restriction endonuclease sites present in the 3.9-kilobase *Eco*RI fragment and new restriction sites generated by various Tn*5* insertions. The region of homology to *K. pneumoniae nifD* was localized to a 1.5-kilobase region between Tn*5* insertions 2 and 10. The region of homology to *K. pneumoniae nifH* was localized to a 1.0-kilobase region to the right of Tn*5* insertion 10 and to the left of *Xho*I site B on pRmR2. R, *Eco*RI restriction site; H, *Hin*dIII restriction site; X, *Xho*I restriction site; B, *Bam*HI restriction site; Bg, *Bgl*III restriction site.

circumvent the tedium of this approach, we have developed new physical and genetic methods to identify and manipulate *R. meliloti nif* genes.

We previously showed that the *K. pneumoniae* DNA sequences coding for one subunit of nitrogenase (*nifD*) and nitrogenase reductase (*nifH*) share DNA homology with 13 widely divergent nitrogen-fixing species, including *R. meliloti*[1]. We used this homology to clone a 3.9-kilobase *R. meliloti Eco*RI fragment in the plasmid cloning vector pACYC184 (ref. 3) (pRmR2; see Fig. 1). To obtain genetic verification that pRmR2 contains essential *nif* genes and to initiate a detailed genetic analysis of *R. meliloti nif* genes, we developed a method to replace the *R. meliloti* wild-type genomic *nif* genes with homologous sequences altered by transposon Tn*5* mutagenesis. We chose transposon Tn*5* for this mutagenesis because it confers resistance to kanamycin (Km) and neomycin (Nm) in *E. coli* and *R. meliloti*, exhibits little insertion site specificity and in general causes polar insertion mutations[4,5]. The experimental strategy we adopted involved the series of steps summarized below and in Fig. 2.

Step I: Plasmid pRmR2 was mutagenized with Tn*5* by infecting an *E. coli* strain containing pRmR2 with λ::Tn*5* (ref. 6). Plasmids carrying Tn*5* insertions were isolated and the locations of the Tn*5* insertions determined by restriction enzyme analysis (see Fig. 2 legend). Six plasmids carrying Tn*5* insertions at different locations along the entire length of the 3.9-kilobase fragment were selected for further study (see Fig. 1). We will refer to these as '*nif* region::Tn*5*' insertions.

Step II: A DNA transformation system has not been developed for *R. meliloti*. Therefore, to introduce the *nif* region::Tn*5* insertions into *R. meliloti*, each *nif* region::Tn*5* *Eco*RI fragment was recloned into the conjugative cloning vector pRK290. (Plasmid pRK290, constructed by Ditta *et al.*[6], is a 20-kilobase derivative of the P-group conjugative plasmid RK2, confers tetracycline (Tc) resistance and contains a single *Eco*RI site suitable for cloning.) The resulting pRK290-*nif* region::Tn*5* hybrid plasmids were conjugated into *R. meliloti*. Total DNA was isolated from the *R. meliloti* exconjugants and analysed by the Southern gel transfer and hybridization procedure[7]. The pRK290-*nif* region::Tn*5* hybrid plasmids had the expected restriction map and replicated stably in *R. meliloti* (see Fig. 3).

Step III: Our strategy for isolating *R. meliloti* strains in which the plasmid-borne *nif* region::Tn*5* fragment had replaced the genomic wild-type *nif* region was based on the assumption that homologous recombination between pRK290-*nif* region::Tn*5* and the *R. meliloti* genome would occur in a small subpopulation of cells and result in the 'transfer' of the Tn*5* insertion mutation from the plasmid to the genome. Cells in which this recombination event had occurred were detected by conjugation of a second P-group plasmid, pR751 (ref. 8), which is incompatible with pRK290, into the *R. meliloti* strains containing the pRK290-*nif* region::Tn*5* plasmid, and simultaneous selection for neomycin resistance (retention of Tn*5*) and trimethoprim (Tp) resistance (conferred by pR751). All the Tp^rNm^r exconjugants were Tc^s, indicating loss of pRK290. These *R. meliloti* $Tp^rNm^rTc^s$ exconjugants could have resulted from Tn*5* transpositions to pR751 or to the *R. meliloti* genome as well as from homologous recombination between pRK290-*nif* region::Tn*5* and the *R. meliloti* genome. To demonstrate that Tn*5* had recombined into the *R. meliloti* genome by means of flanking *nif* DNA sequence homology, DNA was isolated from the $Tp^rNm^rTc^s$ exconjugants, cleaved with *Bgl*II and analysed by the Southern gel transfer and hybridization technique using

Table 1 Level of nitrogenase activity in whole plants with nodules containing various *Rhizobium meliloti* strains

Strain	Ethylene produced after 24 h in acetylene gas (nmol ml^{-1})
Alfalfa plant alone	<1
+Rm 1021	500 ± 100
+Rm 1021/pRK290	400 ± 100
+Rm 1021/pR751	400 ± 100
+Rm 1021/pRK290-*nif* region::Tn*5*	500 ± 100
+Rm 1021 *nif* region::Tn*5* no. 2	NT
+Rm 1021 *nif* region::Tn*5* no. 12	<1
+Rm 1021 *nif* region::Tn*5* no. 7	<1
+Rm 1021 *nif* region::Tn*5* no. 10	<1
+Rm 1021 *nif* region::Tn*5* no. 22	500 ± 100
+Rm 1021 *nif* region::Tn*5* no. 20	500 ± 100
+Rm 1021 *met*::Tn*5*	300 ± 100

Alfalfa seeds (variety Iroquois) were sterilized for 30 min in 70% ethanol followed by 30 min in 50% commercial bleach, and then germinated in the dark in distilled water for 3 days. Individual seedlings were planted on 6-ml slants of nitrogen-free agar (0.1% $CaHPO_4$, 0.02% K_2HPO_4, 0.02% $MgSO_4 \cdot 7H_2O$, 0.02% NaCl, 0.01% $FeCl_3$, 1.0% agar)[13] in 18 × 150 mm test tubes and grown at 27 °C with overhead illumination of 200 foot-candles on 16-h days for 7 days before bacterial inoculation. A freshly grown colony of *R. meliloti* grown on LB agar plus appropriate drugs was picked with a flat toothpick, suspended in 5 ml distilled water and used to inoculate five alfalfa seedlings. After 4 weeks of growth in the conditions described above, the tubes were sealed with serum stoppers, injected with 1 ml acetylene gas per tube, and incubated at room temperature in the dark for 24 h. Ethylene gas production was monitored by gas chromatography as described previously[14]. Control plants not inoculated with bacteria did not form nodules and no acetylene reduction was detected. All plants inoculated with *R. meliloti* formed nodules but those inoculated with strains containing nif^-::Tn*5* insertions did not exhibit detectable levels of nitrogenase activity and had approximately 5–10 times the normal number of nodules per plant. Both the control plants and the plants inoculated with Nif^- bacterial strains were stunted and yellow after 4 weeks of growth compared with plants inoculated with Nif^+ bacterial strains, which were green and visibly larger. NT, not tested.

^{32}P-labelled pRmR2 DNA as a hybridization probe. Homologous recombination of each *nif* region :: Tn*5* insertion from the plasmid into the *R. meliloti* genome *nif* region would result in a different predictable change in the restriction map of that region whereas transposition of Tn*5* would not be expected to change the restriction map of the genomic *nif* region. Figure 3*a*, *b*, *c* illustrates the expected *Bgl*II restriction fragments homologous to pRmR2 for various strains intermediate in the transfer of Tn*5* from pRK290-*nif* region :: Tn*5* to the *R. meliloti* genome. Figure 3*d* is an autoradiogram of an actual hybridization and shows that in the Tp^{r}Nm^{r}Tc^s strains examined, each *nif* region :: Tn*5* had replaced the normal 3.9-kilobase *nif* region in the *R. meliloti* genome through a homologous recombination event. Among 24 Tp^{r}Nm^{r}Tc^s colonies analysed, none contained Tn*5* transpositions to new locations in the genome or to pR751, and none contained replicon fusions between pR751 and pRK290-*nif* region :: Tn*5* (ref. 9).

Step IV: Each *R. meliloti* strain containing a verified Tn*5* insertion in the *nif* region of the genome was inoculated onto a sterile alfalfa seedling and nitrogenase activity was determined 3–4 weeks later using the acetylene reduction technique as described in Table 1 legend. Among five strains containing Tn*5* insertions in the *nif* region, three had undetectable levels of nitrogenase activity (Nif$^-$) on establishment of symbiosis, whereas two had normal nitrogenase activity (Nif$^+$) (Fig. 1, Table 1). The construction of two Nif$^+$ strains with genomic Tn*5* insertions in the 3.9-kilobase *Eco*RI fragment demonstrates that the construction of the three Nif$^-$ :: Tn*5* mutations in unlikely to be an artefact of the cloning and recombinational techniques used in our gene replacement strategy. All the strains with a normal genomic *nif* region, and also plasmid pRK290-*nif* region :: Tn*5*, were Nif$^+$, indicating that none of the *nif* region :: Tn*5* insertions was a dominant *nif*$^-$ mutation. All five strains containing Tn*5* recombined into the genome were prototrophic, indicating that the Nif$^-$ phenotype of three of the strains was not an indirect effect of auxotrophy.

In conjunction with the data presented in Table 1 and Fig. 3, which correlate Tn*5* map position with Nif phenotype, we have used pRmR2 :: Tn*5* plasmids to map the approximate positions of *R. meliloti nif* genes on the 3.9-kilobase *Eco*RI fragment. In these experiments we used specific *K. pneumoniae* restriction fragments carrying regions of genes *nifD* (nitrogenase) and *nifH* (nitrogenase reductase) as hybridization probes to various restriction digests of pRmR2 and pRmR2 :: Tn*5* plasmids (Fig. 1). We found that sequences within both *K. pneumoniae nifD* and *nifH* are conserved in *R. meliloti*. The conserved *R. meliloti* '*nifD*' sequences are localized to 1.5 kilobases between Tn*5* insertions 2 and 10, and the conserved *R. meliloti* '*nifH*' sequences are localized to 1.0 kilobase between Tn*5* insertion 10 and *Xho*I site B indicated in Fig. 1.

As expected on the basis of the interspecies DNA homology with *K. pneumoniae nif* genes, those Tn*5* insertions which cause a Nif$^-$ phenotype in *R. meliloti* are localized near or in these regions of homology, whereas the two Tn*5* insertions which have no effect on *nif* gene expression are located at the other end of the 3.9-kilobase *R. meliloti Eco*RI fragment and not in the region of interspecies *nif* gene homology. Because Tn*5* insertion 22 does not cause a Nif$^-$ phenotype, the *nif* gene *H* equivalent in *R. meliloti* probably starts to the left of this Tn*5* insertion (Fig. 1). The protein subunits for nitrogenase and nitrogenase reductase have average molecular weights of 55,000 and 28,000, respectively, in other *Rhizobium* species[10,11]; therefore, the minimum length of DNA necessary to code for these proteins should be 2.2 kilobases. Thus, it is likely that the 2.6 kilobases between Tn*5* insertion 22 and Tn*5* insertion 2 contains all of *nifH* and at least part of *nifD*.

An important conclusion to be drawn from the present data is that cloned *R. meliloti* DNA sequences homologous to *K. pneumoniae nifD* and *nifH* are essential for symbiotic nitrogen fixation. The data rule out the possibility that the DNA homology between *K. pneumoniae nif* genes and *R. meliloti* DNA is due to cryptic *nif* genes not functional in symbiotic

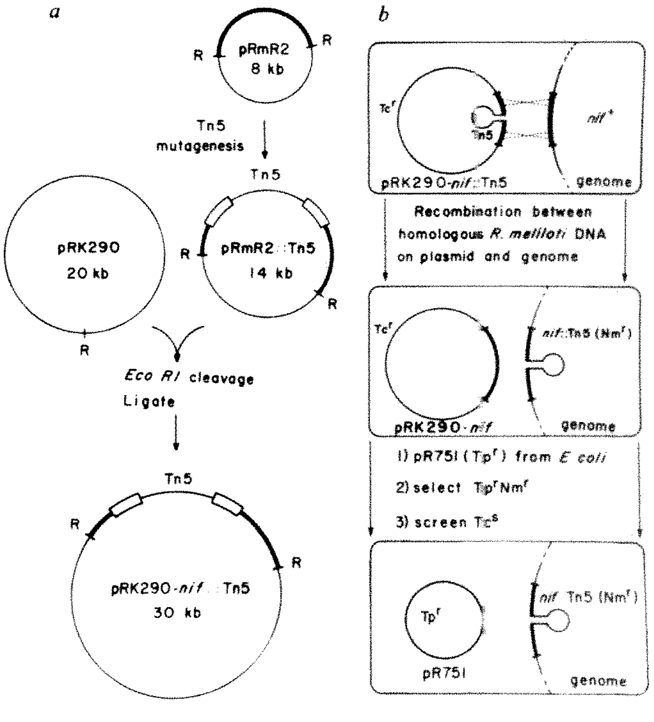


Fig. 2 Outline of experimental strategy used to introduce specific Tn*5* mutations into the *R. meliloti* genome. *a*, Mutagenesis of pRmR2 with Tn*5* and recloning of *nif* region :: Tn*5* fragments into pRK290. *E. coli* strain HB101/pRmR2 (*thr leu thi recA hsdR hsdM pro Str^r* Tc^r) was grown at 37 °C in LB medium containing 0.2% maltose and 10 μg ml^{-1} Tc to a density of 5×10^8 cells per ml and infected with λ*b*221 *rex* :: Tn*5 cI*857 O_{am8} P_{am29} (ref. 5) at a multiplicity of 0.1 particles per cell. Approximately 5×10^3 kanamycin-resistant survivors were obtained per plate when 0.2 ml of infected cells were plated on LB agar containing 20 μg ml^{-1} Km at 32 °C. Colonies from three plates were resuspended in 10 ml of 10 mM $MgSO_4$, plasmid DNA was isolated[15] and used to transform $CaCl_2$-treated strain HB101 (ref. 16), and Tc^r and Tc^{r}Km^r transformants were selected on LB agar containing 10 μg ml^{-1} Tc and 20 μg ml^{-1} Km. Approximately 20 Tc^{r}Km^r transformants were obtained per 10^4 Tc^r transformants. Plasmid DNA was prepared from 26 Tc^{r}Km^r colonies using a small-scale purification procedure[17], and the plasmid DNA was digested in UREB with *Eco*RI. Because Tn*5* is 5.7 kilobases and does not contain an *Eco*RI site, insertions of Tn*5* into the 3.9-kilobase *nif* homologous *R. meliloti Eco*RI restriction fragment were monitored (using agarose-gel electrophoresis) by the appearance of a 9.6-kilobase *Eco*RI restriction fragment and disappearance of the 3.9-kilobase *Eco*RI restriction fragment. Among the 26 plasmids, 11 contained Tn*5* insertions in the 3.9-kilobase *Eco*RI fragment and the sites of these Tn*5* insertions along the 3.9-kilobase *Eco*RI restriction fragment were determined by double restriction endonuclease digestions of each pRmR2 :: Tn*5* plasmid in 1X UREB with *Xho*I and *Eco*RI or *Hin*dIII and *Eco*RI. Six pRmR2 :: Tn*5* plasmids with Tn*5* insertions distributed along the entire length of the 3.9-kilobase *Eco*RI fragment were chosen for further analysis. A 0.5-μg aliquot of each of the six pRmR2 :: Tn*5* plasmids was digested with *Eco*RI as described above and mixed with a 0.5-μg aliquot of pRK290 DNA which had also been digested with *Eco*RI. The mixed DNAs were ligated with 1 unit of T4 DNA ligase, as described elsewhere[14], and the ligated DNA was used to transform *E. coli* strain MM294 (*endo*I *hsd*R *pro*). Following a nonselective growth in 1 ml LB for 2 h at 32 °C, a 0.1-ml aliquot was diluted into 5 ml LB containing 10 μg ml^{-1} Tc and 20 μg ml^{-1} Km and grown to saturation at 32 °C. 0.1 ml of this culture was mated simultaneously with 0.1 ml *R. meliloti* strain Rm1021 (nif$^+$Str^r) and 0.1 ml MM294/pRK2013, and incubated on an LB agar plate overnight at 32 °C. (Plasmid pRK2013 (ref. 18) contains the *tra* genes of RK2 cloned in plasmid ColE1 and complements pRK290 *in trans* for conjugation of pRK290.) Str^{r}Tc^{r}Nm^r *R. meliloti* exconjugants were selected on LB agar containing 250 μg ml^{-1} streptomycin (Str), 5 μg ml^{-1} Tc and 20 μg ml^{-1} Nm. Because pRK2013 (which confers Nm^r) cannot replicate in *R. meliloti*[6], this selection yielded *R. meliloti* exconjugants containing pRK290-*nif* region :: Tn*5* plasmids which arose at a frequency of 10^{-3}. *b*, Selection of *R. meliloti* strains in which the *nif* region :: Tn*5* mutations had recombined in the genome. *E. coli* strain J53/pR751 (ref. 8) was conjugated with each of the *R. meliloti* M1021/pRK290-*nif* region :: Tn*5* strains as described above and Str^{r}Tp^{r}Nm^r exconjugants were selected on LB agar containing 250 μg ml^{-1} Str, 700 μg ml^{-1} Tp and 30 μg ml^{-1} Nm. Exconjugants arose at a frequency of 10^{-6} and 24 of these (four from each of the six pRK290-*nif* region :: Tn*5* strains) were analysed as described in Fig. 3 legend. *R. meliloti nif* region DNA is shown in heavy lines; Tn*5* inverted repeat DNA shown as open boxes (R = *Eco*RI restriction site).

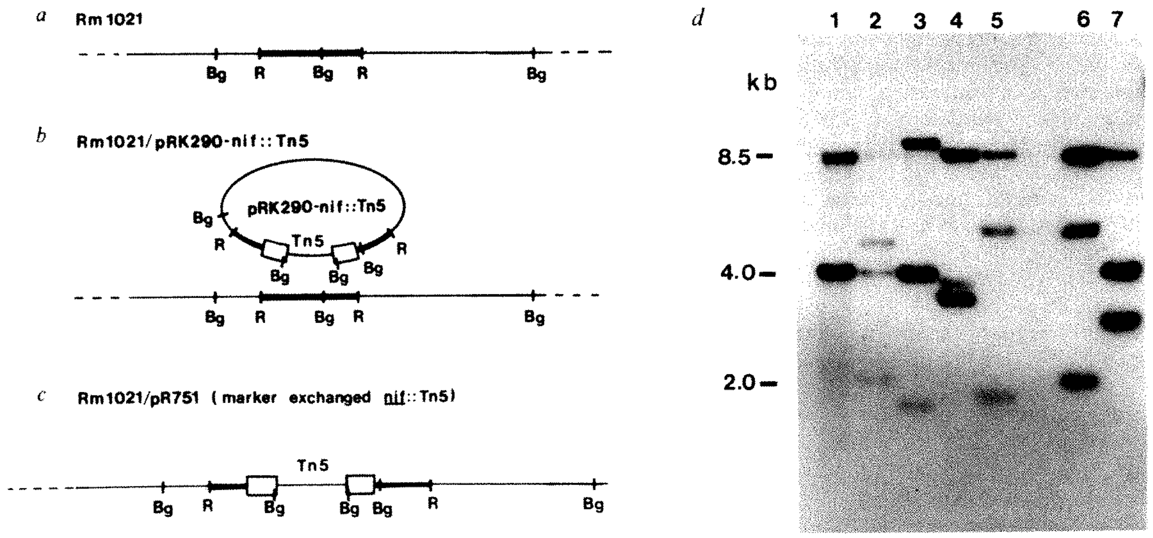


Fig. 3 Change in the *Rhizobium meliloti nif* region restriction map due to homologous recombination of *nif* region :: Tn*5* from pRK290 to the *R. meliloti* genome. A restriction map of the *nif* region in the *R. meliloti* genome was constructed as follows. Total DNA was isolated from *R. meliloti* strain Rm1021, using the Marmur technique[19]; approximately 1 μg of DNA was digested with restriction endonuclease *Bgl*II or *Eco*RI (5 U Bethesda Research Laboratories) or both, and restriction fragments were separated on 0.7% agarose gels and transferred to nitrocellulose by the method of Southern[7] as described elsewhere[2]. Plasmid pRmR2 was isolated by the cleared lysate technique[15] and labelled by nick translation[20,21] with [α-^{32}P]TTP (350 Ci mmol^{-1}, Amersham) and DNA polymerase I (Boehringer Mannheim) to a specific activity of 5×10^7 c.p.m. per μg of DNA, and 10^6 c.p.m. of labelled DNA probe was hybridized to each 10 × 15-cm nitrocellulose filter and exposed to Kodak XR-5 film as described elsewhere[2]. *a*, Restriction map of the normal *nif* region of *R. meliloti*. Hybridization of pRmR2 probe to *Bgl*II-digested Rm1021 DNA as described above should yield hybridization bands at 8.5 and 4.0 kilobases which overlap the 3.9-kilobase *Eco*RI fragment (shown in heavy lines) cloned in pRmR2. (The cloning vector pACYC184 does not hybridize to *R. meliloti* DNA; unpublished results.) *b*, Restriction map of the two regions of homology to pRmR2 in Rm1021 strains containing plasmid pRK290-*nif* region :: Tn*5*. Tn*5* insertion into any *Bgl*II restriction fragment adds two *Bgl*II sites each located 1.5 kilobases from the Tn*5* insertion site. Hybridization of pRmR2 probe to *Bgl*II-digested DNA from a Rm1021/pRK290-*nif* region :: Tn*5* strain as described above should yield the normal genomic 8.5- and 4.0-kilobases hybridization bands plus three additional hybridization bands due to the four *Bgl*II sites in each pRK290-*nif* region :: Tn*5* plasmid. *c*, Restriction map of the *nif* region of *R. meliloti* following homologous recombination of Tn*5* into the parent 3.9-kilobase *Eco*RI restriction fragment. Hybridization of ^{32}P-labelled pRmR2 DNA to *Bgl*II-digested DNA from a strain containing a genomic Tn*5* insertion into the parental *Eco*RI fragment cloned in pRmR2 should yield three bands, one of which should be either 4.0 or 8.5 kilobases (same as wild type) and two which should vary in size. The sum of the lengths of these latter two bands should be either 8.5 kilobases or 4.0 plus 3.0 kilobases (2 × 1.5 kilobases of Tn*5* DNA). Because each Tn*5* insertion which was used in the marker-exchange procedure was previously mapped on the parent pRmR2 :: Tn*5* plasmid, a unique *Bgl*II restriction hybridization map for the *nif* region could be predicted for each genomic Tn*5* insertion. For example, Tn*5* insertion 12 is located to the 'left' of the *Bgl*II site on pRmR2 as drawn here. As predicted, hybridization of ^{32}P-labelled pRmR2 DNA to *Bgl*II-digested DNA from a strain containing Tn*5* insertion 12 recombined into the *R. meliloti* genome yielded hybridization bands of 8.5, 3.8 and 3.4 kilobases (*d*, lane 4). Each of the other insertion mutations also gave the predicted hybridization patterns (*d*). *d*, Autoradiogram of ^{32}P-labelled pRmR2 hybridized to DNA from various *R. meliloti* strains (1 μg), digested with *Bgl*II (5 U, in 1X UREB), separated by 1.0% agarose-gel electrophoresis and transferred to nitrocellulose as described above from the following strains: lane 1: Rm1021/pR751 (hybridization bands at 8.5 and 4.0 kilobases); lane 2: Rm1021/pRK290-*nif* region :: Tn*5* no. 10 (hybridization bands at 20, 8.5, 4.8, 4.0 and 2.0 kilobases); lane 3: Rm1021 *nif* region :: Tn*5* no. 22/pR751 (hybridization bands at 9.8, 4.0 and 1.7 kilobases); lane 4: Rm1021 *nif* region :: Tn*5* no. 12/pR751 (hybridization bands at 8.5, 3.7 and 3.3 kilobases); lane 5: Rm1021 *nif* region :: Tn*5* no. 7/pR751 (hybridization bands at 8.5, 5.2 and 1.8 kilobases); lane 6: Rm1021 *nif* region :: Tn*5* no. 10/pR751 (hybridization bands at 8.5, 5.1 and 1.9 kilobases); lane 7: Rm1021 *nif* region :: Tn*5* no. 20/pR751 (hybridization bands at 8.5, 4.0 and 2.9 kilobases). *Bgl*II-digested DNA from strains M1021/pRK290, M1021/pR751 and M1021 all hybridized equivalently to pRmR2 probe (data not shown). The lower portion of the autoradiograph was photographically enhanced to show the faint smaller hybridization bands. Bg, *Bgl*II restriction site; R, *Eco*RI restriction site; kb, kilobases.

nitrogen fixation or due to homology with genes evolutionarily (or fortuitously) related to *K. pneumoniae nif* genes. Based on the interspecies homology with *K. pneumoniae nifD* and *nifH* and on the mutational analysis described here, we conclude that *R. meliloti* contains genes equivalent to *K. pneumoniae nifD* and *nifH*, although we cannot yet assign these homologous *R. meliloti nif* genes to separate complementation groups. Because we expect other nitrogen-fixing species to have similar genes, we propose the following nomenclature for nitrogenase genes from various organisms. Using the letters already assigned to *nif* genes in *K. pneumoniae*, we suggest adding the first letters of the genus and species names. Thus *K. pneumoniae nifH* would be written Kp *nifH* and the corresponding gene in *R. meliloti* would be written Rm *nifH*. This nomenclature is analogous to that already in use for distinguishing nitrogenase proteins isolated from different species[12].

Because the techniques described use a wide host-range conjugative cloning vector and do not necessarily require previous identification of genetic function of the cloned DNA, they should be applicable to fine-structure genetic analysis of a wide variety of Gram-negative prokaryotes. These techniques should also be useful in the initiation of genetic studies of DNA fragments cloned from Gram-negative prokaryotes by non-genetic criteria, for example, homology to inducible RNA or homology to specific DNA sequences. Finally, DNA fragments identified and cloned on the basis of transposon insertional inactivation can be further mutagenized using these techniques.

We thank G. Ditta for providing plasmids pRK290 and pRK2013 before publication, G. Jacoby for providing plasmid pR751, F. Lang for advice on decreasing background in Tp selections, and R. Hyde for preparation of this manuscript. This work was supported by USDA/SEA/CRGO grant 7900481 and NSF grant PCM7806834 awarded to F. M. A.

Received 21 July; accepted 16 October 1980.

1. Ruvkun, G. B. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 191–195 (1980).
2. Vincent, J. M. in *The Biology of Nitrogen Fixation* (ed. Quispel, A.) 265–341 (North-Holland, Amsterdam, 1974).
3. Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134,** 1141–1156 (1978).
4. Jorgenson, R. A., Rothstein, S. J. & Reznikoff, W. S. *Molec. gen. Genet.* **177,** 65–72 (1979).
5. Berg, D. E. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, T. A. & Adhya, S. L.) 205–212 (Cold Spring Harbor Laboratory, New York, 1977).
6. Ditta, G., Stanfield, S. Corbin, D. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
7. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
8. Jobanputra, R. S. & Datta, N. *J. med. Microbiol.* **7,** 169–177 (1974).
9. Kleckner, N. *Cell* **11,** 11–23 (1977).
10. Isreal, D. W., Howard, R. L., Evans, H. J. & Russell, S. A, *J. biol. Chem.* **249,** 500–508 (1974).
11. Whiting, M. J. & Dilworth, M. J. *Biochim. biophys. Acta* **371,** 337–351 (1974).
12. Burns, R. C. & Hardy, R. W. F. *Nitrogen Fixation in Bacteria and Higher Plants,* 64 (Springer, New York, 1975).
13. Vincent, J. M. in *A Manual for the Practical Study of Root-Nodule Bacteria,* 75 (Blackwell, Oxford, 1970).
14. Cannon, F. C., Riedel, G. E. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2963–2967 (1977).
15. Clewell, D. B. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **62,** 1159–1166 (1969).
16. Cohen, S. N., Chang, A. C. Y. & Hsu, C. L. *Proc. natn. Acad. Sci. U.S.A.* **69,** 2110–2124 (1972)
17. Klein, R. D., Selsing, E. & Wells, R. D. *Plasmid* **3,** 88–91 (1980).
18. Figurski, D. H. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1648–1652 (1979).
19. Marmur, J. *J. molec. Biol.* **3,** 208–218 (1961).
20. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113,** 237–251 (1977).
21. Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
22. Cannon, F. C., Riedel, G. E. & Ausubel, F. M. *Molec. gen. Genet.* **174,** 59–66 (1979).
23. Riedel, G. E., Ausubel, F. M. & Cannon, F. C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2866–2870 (1979).

Escherichia coli ribosomal protein S8 feedback regulates part of *spc* operon

Dennis Dean, John L. Yates & Masayasu Nomura*

Institute for Enzyme Research, Departments of Genetics and Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706

In *Escherichia coli* the genes coding for the 52 ribosomal proteins (r-proteins) are organized into a number of transcription units located at various regions on the bacterial genome. The expression of r-protein genes is balanced so that individual r-protein synthesis rates change coordinately in response to changing environmental conditions, and significant amounts of free r-proteins do not exist in the cellular pool. We have suggested a model for the balanced regulation of r-protein gene expression, namely that r-protein synthesis and ribosome assembly are coupled so that r-proteins not incorporated into ribosomes prevent the further translation of r-protein mRNA by a feedback regulatory mechanism[1]. The model was tested *in vitro* by examining the effect of purified r-proteins on DNA directed r-protein synthesis[2,3], and *in vivo* by examining the effect of overproduction of certain r-proteins on the synthesis rates of other r-proteins[4,5]. *In vitro* experiments have revealed that some r-proteins (L1, L4, L10, S4 and S8) can selectively inhibit the synthesis of r-proteins whose genes are in the same operon as their own, and that this specific feedback regulation occurs at the level of translation rather than at the level of transcription of mRNA[2,3,6,7]. Regulatory roles for L1, S4 and L4 have also been established by *in vivo* experiments[4,5]. We have studied further the feedback regulatory properties of S8 *in vivo* and *in vitro*, and report here that the protein regulates a part of the *spc* operon.

Figure 1 shows the *in vitro* effects of purified S8 addition on synthesis of r-proteins encoded in the *spc* operon. Protein synthesis was directed by a DNA fragment containing most genes of the *spc* operon (L14, L24, L5, S14, S8, L6 and L18) or by λ*spc*1 DNA. λ*spc*1 transducing phage carries the genes in the *spc* and α operons (see Fig. 2). Radioactive proteins were precipitated from reaction mixtures with specific antisera and analysed by SDS-gel electrophoresis followed by autoradiography. Addition of S8 (2.9 μM) results in specific inhibition of L5 and S14 synthesis (Fig. 1*a*, 5–8), but only slight or no inhibition of other proteins from the *spc* operon. Densitometric tracings of films from several such experiments showed that S8 inhibits the synthesis of L5 and S14 by 75–90% but inhibits the synthesis of L14, L24, S8, L6, L18 and S5 by less than 30%. The effect of S8 addition on the synthesis of L30 and L15 was not examined.

We have previously reported that L24 synthesis was also inhibited by S8 (ref. 2). However, we have now established that S8 has no significant inhibitory effect on L24 synthesis in the DNA-directed system using immunoprecipitation methods and consider that the inhibition originally observed was inhibition of synthesis of S14, which co-migrates with L24 in the SDS gels used. As we have previously shown[2], addition of other r-proteins such as L1 or S4 has no inhibitory effects on synthesis of r-proteins contained within the *spc* operon. It is significant that expression from the promoter proximal genes (L14 and L24) is not inhibited by S8. These results support our previous conclusion that regulatory r-proteins act at the level of translation and not at the level of transcription originating from the promoter.

To test whether the regulatory effects of S8 observed *in vitro* also reflect *in vivo* regulatory properties of S8, we examined the effect of S8 over-production on synthesis rates of other r-proteins *in vivo*. A DNA fragment containing three genes from the *spc* operon (S8, L6 and L18) was cloned into a plasmid vector containing the *lac* operator and promoter so that the synthesis of S8, L6 and L18 were then controlled by the *lac* regulatory elements (pNO1017; see Fig. 2). Another plasmid (pNO1018) that contains only the S8 gene fused to the *lac* operator–promoter was derived from pNO1017 (see Fig. 2).

Effects of over-production of S8, L6 and L18, or S8 alone on the synthesis rates of other individual r-proteins were examined by stimulating transcription from *lacZp* in strains which harbour the recombinant plasmids carrying r-protein genes fused to *lacZp*. Exponentially growing cells were treated with an inducer of the *lac* operon, 1 mM isopropyl thiogalactoside (IPTG), to stimulate synthesis of the various cloned r-protein genes. Synthesis rates of individual r-proteins were then analysed by labelling cells with [^{3}H]lysine for 1 min followed by a 1-min chase with nonradioactive lysine. Relative synthesis rates were obtained by comparing the synthesis rates of individual r-proteins determined after IPTG induction with the synthesis rates of individual r-proteins before IPTG addition.

Table 1 Relative synthesis rates of ribosomal proteins after induction by IPTG

	Strains	
r-protein	NO2320 (carries pNO1017)	NO2321 (carries pNO1018)
S4	1.21	1.14
S5	0.50*	0.25*
S7	1.07	1.10
S8	4.10*	2.25*
S13	1.13	1.12
S14	0.37*	0.37*
S20	1.21	1.13
L1	$\boxed{1.00}$	$\boxed{1.00}$
L2	1.13	1.13
L3	1.00	1.00
L5	0.36*	0.36*
L6	3.50*	0.40*
L13	0.99	0.97
L14	1.14	1.02
L15	0.47*	0.57*
L16	1.02	1.05
L17	1.02	1.13
L18	3.55*	0.31*
L19	1.10	0.99
L22	1.24	1.22
L24	0.99	1.04
L28	1.23	1.04
L30	0.26*	0.30*

For labelling experiments cells were grown in a synthetic minimal medium[12] supplemented with 0.4% glycerol, all amino acids (except methionine and lysine) at 50 μg ml^{-1}, and thiamine at 2 μg ml^{-1}. Cells were grown at 37 °C to 2×10^8 ml^{-1}. Samples of cells were pulse-labelled with ^{3}H-lysine (646 pmol ml^{-1}; 80.5 Ci mmol^{-1}) for 1 min followed by a 1 min 'chase' with excess nonradioactive lysine. Cells were labelled before (control cells), and 10 min after addition of IPTG to 1 mM (experimental cells). Cells were rapidly chilled and then mixed with a suitable amount of ^{14}C-lysine-labelled carrier cells. ^{14}C-labelled cells were prepared by growing cells in the presence of ^{14}C-lysine for several generations. r-Proteins were extracted and separated by two-dimensional gel electrophoresis[13]. After the gels had been stained, the spots corresponding to the indicated r-proteins were extracted, the gels were oxidized with a Packard sample oxidizer, and ^{3}H and ^{14}C measured separately. The ^{3}H/^{14}C ratio in each protein in the experimental cells was then compared with that in control cells using the expression:

$$\frac{[^3\mathrm{H}/^{14}\mathrm{C}]\text{ in protein } i \text{ in experimental cells}}{[^3\mathrm{H}/^{14}\mathrm{C}]\text{ in protein } i \text{ in control cells}}$$

The ratios obtained were then normalized to the ratio obtained for L1 (indicated by a box). Values given in the table are the average of two independent experiments. *Values significantly different from 1.0.

---

* To whom correspondence should be addressed.

0028-0836/81/020089—03$01.00

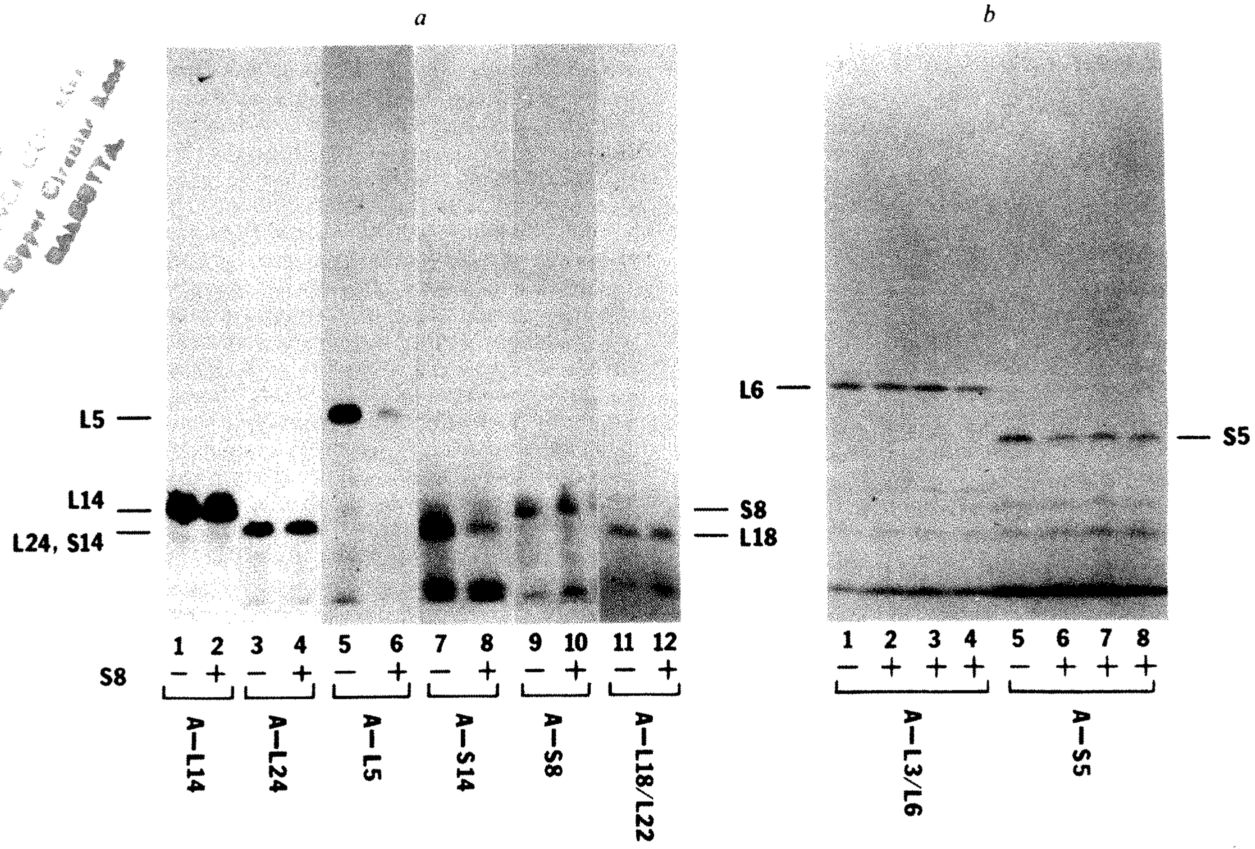


Fig. 1 Effects of S8 on protein synthesis from the *spc* operon *in vitro*. *a*, *In vitro* protein synthesis was carried out as described elsewhere[2] using the 10% *Eco*RI DNA fragment of λ*spc*2 transducing phage as template. This fragment carries genes between an *Eco*RI site in the S3 gene from the S10 operon and the *Eco*RI site in the S5 gene from the *spc* operon (see Fig. 2). Protein was labelled using ^{35}S-methionine (160 mCi per μmol, 15 μM). Two reaction mixtures (80 μl) received either no S8 (control sample) or 3 μg S8. The reactions were terminated after 60 min at 37 °C, then 3 μg S8 was added to the control sample. 10-μl aliquots (5 μl for 1 and 2) were taken from reaction mixtures which had been incubated in the absence or presence of S8, as indicated, and radioactive proteins synthesized were precipitated as described previously[2] by treating with antisera to specific r-proteins. The antiserum used for a particular experiment is shown at the bottom of the figure (A followed by the r-proteins used to raise that antiserum). A-L18/L22 is an antiserum raised against a mixture of r-proteins L18 and L22. Polyacrylamide gel electrophoresis was carried out according to the procedure of Laemmli[8] in 13.5% acrylamide-SDS gels. Gels were impregnated with diphenyloxazole (PPO)[9] and exposed to pre-flashed X-ray film[10]. In the autoradiograms shown the direction of migration is from the top to the bottom. The origin of radioactive protein bands migrating to the bottom of the gel is not known. No consistent pattern in the amounts of this material was observed in relation to the presence or absence of S8 in the reaction mixture. *b*, Proteins were synthesized using λ*spc*1 DNA as template and labelled with ^{3}H-lysine (80 mCi per μmol, 15 μM). Analysis was done as described above using 80 μl reaction mixtures with no S8 added during incubation (lanes 1, 5), 1.8 μg S8 added (lanes 2, 6), 2.7 μg S8 (lanes 3, 7) and 3.6 μg S8 (lanes 4, 8). To analyse L6, an antiserum against a mixture of L3 and L6 (A-L3/L6) was used.

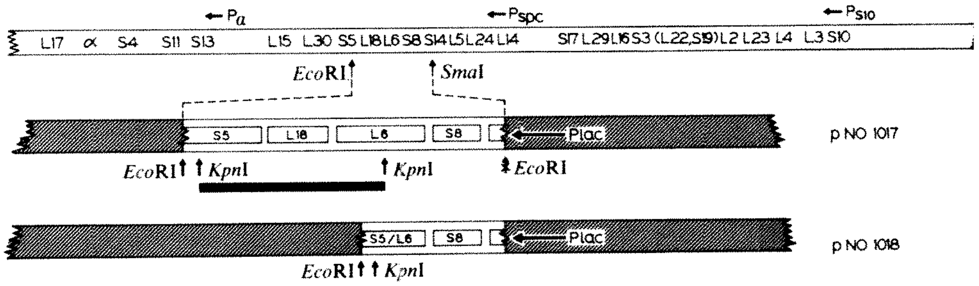

← Pα
← Pspc
← PS10
L17 α S4 S11 S13 L15 L30 S5 L18 L6 S8 S14 L5 L24 L14 S17 L29 L16 S3 (L22,S19) L2 L23 L4 L3 S10
EcoRI
SmaI
S5
L18
L6
S8
Plac
pNO 1017
EcoRI
KpnI
KpnI
EcoRI
S5/L6
S8
Plac
pNO 1018
EcoRI
KpnI

Fig. 2 Construction of hybrid plasmids pNO1017 and pNO1018. The plasmid vector was pBGP120 (ref. 11), which is a derivative of RSF2124 and contains a region of the λ*plac* genome including the *lac* promoter and the main part of the *lacZ* gene. The approximate position of the *lacZp* ('Plac' in the figure) and the direction of transcription from *lacZp* (the thick horizontal arrow) are indicated. The vector region is hatched and the bacterial region is open. Plasmid pNO1017 was constructed by ligating *Eco*RI-digested pBGP120 and the *Eco*RI-*Sma*I fragment indicated in the figure purified from λ*spc*-DNA[2]: the fusion of the *Sma*I-*Eco*RI ends resulted in loss of the *Eco*RI and *Sma*I sites (indicated by a cross through an *Eco*RI site in pNO1017). pNO1018 was constructed by digesting pNO1017 with *Kpn*I and ligation to delete the region indicated by the cross-hatched region. The positions of the *Kpn*I sites are approximate.

Over-production of r-proteins S8, L6 and L18 resulted in an inhibition of synthesis of certain other r-proteins (L5, S14, S5, L30 and L15) contained within the *spc* operon (Table 1). Synthesis rates for the promoter proximal r-proteins (L14 and L24), however, were not affected. Similarly over-production of S8 alone inhibited the synthesis of L5, S14, L6, L18, S5, L30 and L15, but not the synthesis of L14 or L24. No significant changes in the synthesis rates of other r-proteins not contained in the *spc* operon were observed as a consequence of S8, L6 or L18 over-production. The *in vivo* results, therefore, confirm our *in vitro* results—S8 is capable of inhibiting synthesis of certain r-proteins encoded in the *spc* operon. Similar to previous results obtained with the S10 and α operons, synthesis of proteins from promoter distal genes were inhibited *in vivo* but not *in vitro*. As discussed previously[3,5], we believe that the observed regulatory effects on promoter distal genes observed *in vivo* are the consequence of a polar effect caused by direct inhibition of translation of upstream mRNA. The *in vivo* polar effect could be due to premature transcription termination, increased degradation of distal mRNA, or translation of distal cistrons requiring ribosomes that initiated in an earlier cistron. The important observation in the present study is that, unlike the previously studied r-protein operons, inhibition occurs at the third gene in the operon rather than the first. This result clearly indicates that inhibition of synthesis of specific r-proteins contained in the *spc* operon by S8 cannot occur at the level of initiation of transcription from the *spc* promoter. This conclusion is consistent with our previous *in vitro* observations that autogenous feedback regulation by repressor r-proteins occurs at the translation step rather than the transcription step[2,3]. It is particularly intriguing that synthesis of r-proteins that are coordinately regulated and whose genes are contained within a single transcription unit are apparently regulated independently at the translational level. The method of regulation of synthesis of r-proteins L14 and L24 is unknown.

We thank Kristin Ellingsen for technical help. This work was supported in part by the College of Agricultural and Life Sciences, University of Wisconsin-Madison, by grant GM-20427 from the NIH, and by grant PCM 79-10616 from the NSF. D.D. is supported by NIH postdoctoral fellowship grant GM-07553.

Received 29 August; accepted 29 October 1980.

1. Fallon, A. M., Jinks, C. S., Strycharz, G. D. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3411–3415 (1979).
2. Yates, J. L., Arfsten, A. E. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1831–1841 (1980).
3. Yates, J. L. & Nomura, M. *Cell* **21,** 517–522 (1980).
4. Zengel, J. M., Mueckl, D. & Lindahl, L. *Cell* **21,** 523–535 (1980).
5. Dean, D. & Nomura, M. *Proc, natn. Acad. Sci. U.S.A.* **77,** 3590–3594 (1980).
6. Brot, N., Caldwell, P. & Weissbach, H. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2592–2595 (1980).
7. Fukuda, R. *Molec. gen. Genet.* **178,** 483–486 (1980).
8. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
9. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
10. Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56,** 335–341 (1975).
11. Polisky, B., Bishop, R. J. & Gelfand, D. H. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3900–3904 (1976).
12. Clark, D. J. & Maaløe, O. J. *J. molec. Biol.* **23,** 99–112 (1967).
13. Howard, G. A. & Traut, R. R. *FEBS Lett.* **29,** 177–180 (1973).

Oriented fibrin gels formed by polymerization in strong magnetic fields

J. Torbet
Max-Planck-Institut für Festkörperforschung, Hochfeld-Magnetlabor
166X, F-38042 Grenoble Cedex, France

J. -M. Freyssinet & G. Hudry-Clergeon
Laboratoire d'Hématologie, DRF, FRA 21 INSERM, Centre
d'Etudes Nucléaires de Grenoble, 85 X, F-38041 Grenoble Cedex,
France

Fibrinogen is a soluble plasma protein which, after cleavage by the specific proteolytic enzyme thrombin, polymerizes to form the filamentous fibrin network during blood clotting (see refs 1 and 2 for reviews). Fibrinogen has a molecular weight of 340,000 and is composed of two identical halves, each containing three peptide chains designated Aα, Bβ and γ. Fibrin monomers are produced by thrombin which releases the small negatively charged fibrinopeptides A and B. The overall shape of the fibrinogen molecule has not been unequivocally established[1,2]. The trinodular, elongated ($\sim$450 Å long) structure proposed by Hall and Slayter[3] is the most widely accepted model and it has obtained additional support from recent work[4–6]. Fibrin monomers are also about 450 Å long[7] and in fibres they probably have a half-staggered arrangement along the axis[7,8]. The fibres are an assembly of protofibrils whose structure and packing are not reliably known. We report here that highly oriented fibrin gels are formed when polymerization takes place slowly in a strong magnetic field. It is shown that the protofibrils pack into a three-dimensional crystalline lattice. We introduce magnetically induced birefringence as a potential tool for studying polymerization and briefly speculate on the applications of strong magnetic fields.

The sigmoidal variation of the birefringence (Fig. 1), observed when a fibrin gel is formed in a strong magnetic field, can bequalitatively explained as follows. Near the beginning there is little induced birefringence because the fibrinogen molecules and fibrin monomers present have only a small diamagnetic anisotropy[9]. The magnetic orienting energy is much less than the thermal energy kT and consequently they orient poorly. The upsurge in the birefringence commences with polymerization. Now many monomers come together in an ordered way, the diamagnetic anisotropy of the aggregates is many times that of a monomer and the magnetic orienting energy is sufficiently large to allow a high degree of orientation. The value of $\Delta n/C$, where n is the refractive index and C the concentration, at the end of the reaction increases as the rate of polymerization is reduced until a limiting value is reachd which corresponds to the maximum alignment of the fibres in the field used. Thus, when polymerization proceeds too rapidly, the interlinked network develops before the nascent fibres have had time to orient themselves fully.

Scanning micrographs of unoriented and magnetically oriented gels are compared in Fig. 2. The samples suffered distortion while being dried in preparation for electron microscopy. Nevertheless, the orientation of the gel can be easily seen to parallel the field direction.

Figure 3 shows the low-angle neutron diffraction pattern obtained from magnetically oriented fibrin gels. In addition to the familiar[10–12] 223 Å (±5 Å)-spaced meridional reflections, there are equatorial and other Bragg reflections which indicate that the fibrin monomers are arranged in a regular three-dimensional lattice. The sharpness of the meridional reflections, which is due to a fibre axial periodicity, shows that there is a high degree of crystallinity along the fibre axis[10]. These reflections exhibit little arcing and when the gel is tilted slightly from the vertical, the intensity of the higher orders is markedly reduced. Hence, the fibres are highly oriented along the magnetic-field direction. Rotating the gel about the vertical axis does not alter the non-meridional reflections, showing that the samples are rotationally symmetrical and therefore there can be little long-range rotational order between fibres.

The equatorial diffraction arises from the regular, lateral arrangement of the fibrin molecules. In the pattern there is a discrete peak at 184 ± 5 Å and an extended band with a maximum at about $184/5^{1/2}$ Å (Figs 3, 4). The diffraction is not detailed enough to deduce directly the parameters of the unit cell perpendicular to the fibre axis. However, from simple

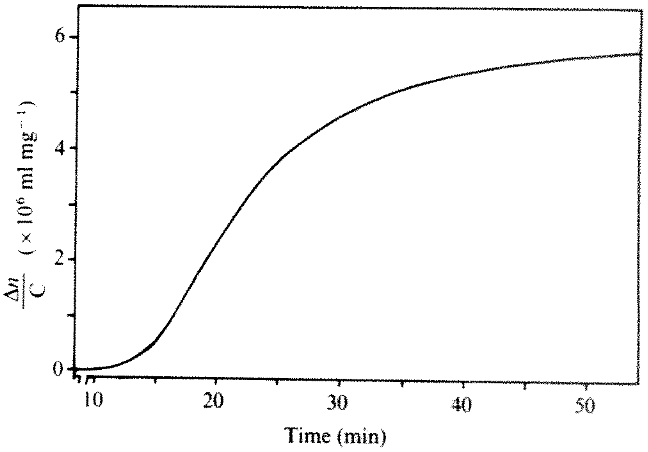


Fig. 1 The change with time of the birefringence per unit concentration of protein $\Delta n/C$ (C in mg ml^{-1}) when polymerization is performed in a constant magnetic field of 11 T (1 T = 10^4 G). The birefringence was measured at 6,328 Å by means of a mechanically driven Babinet Soleil compensator. A small volume of thrombin (0.10 NIH units) was used to induce polymerization at 22 °C of a 1-ml solution of fibrinogen (4.8 mg ml^{-1} in 0.05 M Tris-HCl buffer, pH 7.5, containing 0.10 M NaCl and 0.5 mM EDTA).

028-0836/81/020091—03$01.00

packing considerations one expects the basal unit cell, perpendicular to the fibre axis, to be nearly square or hexagonal. The position of the maximum at $184/5^{1/2}$ Å (Fig. 4) favours a unit cell with a square base of side about 184 Å. As there are few unit cells (23 in a fibre of 1,000 Å diameter) in a fibre cross-section, the relative breadth of the equatorial intensities does not necessarily mean that crystallinity in the lateral plane is less than along the axis. Additionally, fibres of unlike diameters could have slightly different basal unit-cell dimensions which would also give rise to broadening.

Within each quadrant (Fig. 3), between the equator and meridian, there are two maxima. The stronger of these is situated at the intersection of the first layer and row lines, while the second, at a higher angle, has a meridional position between the second and third layer lines, thus the true axial repeat is $n \times 223$ Å where n is an integer. Its smallest possible value is 2, in which case the second off-meridional reflection, if well resolved (its position could be influenced by intensities on adjacent layer lines), should lie halfway between layer lines 2 and 3. Figure 3 shows that this is approximately so but the maximum is broad and weak and consequently its position cannot be determined unequivocally. The value of $n = 2$ is supported by micrographs[7] and agrees with the prevalent view that the approximately 450 Å long molecules form a half-staggered arrangement along the fibre axis[8,10]. Therefore this analysis, which is partly qualified as described above, shows that the diffraction pattern can be indexed on a tetragonal unit cell (with $a = 2 \times 185$ Å and $c = 446$ Å) that contains eight molecules (using 6 $Å^3$ per dalton for the specific volume[13]) and has a space group which restricts reflections to l even when $h = k = 0$. There are no evident restrictions on the general (h, k, l) reflections. None of the four space groups proposed in a recent model building study[13] is supported by the diffraction pattern. However, one of these (space group I4) does have a tetragonal lattice with the unit-cell dimensions we find. Only the space groups $P4_2$ and $P4_222$ satisfy the above conditions.

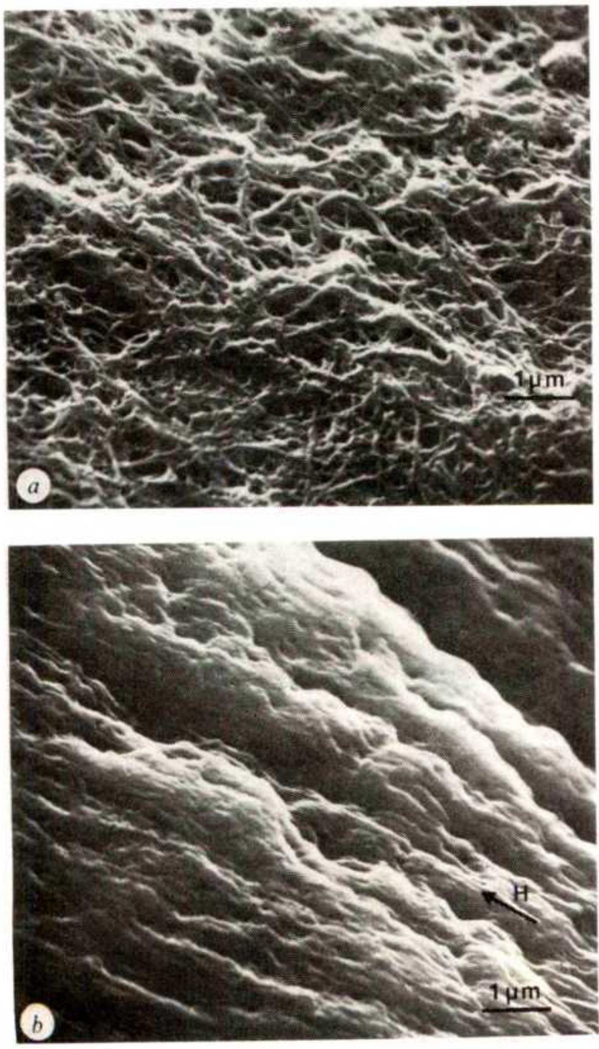


Fig. 2 Scanning electron micrographs of fibrin prepared as described in Fig. 1 legend. *a*, Without field; *b*, with field H in direction of arrow. The diameter of the fibres in *a* is ~1,000 Å. Drying has induced distortion and coalescence of the oriented *b* sample.

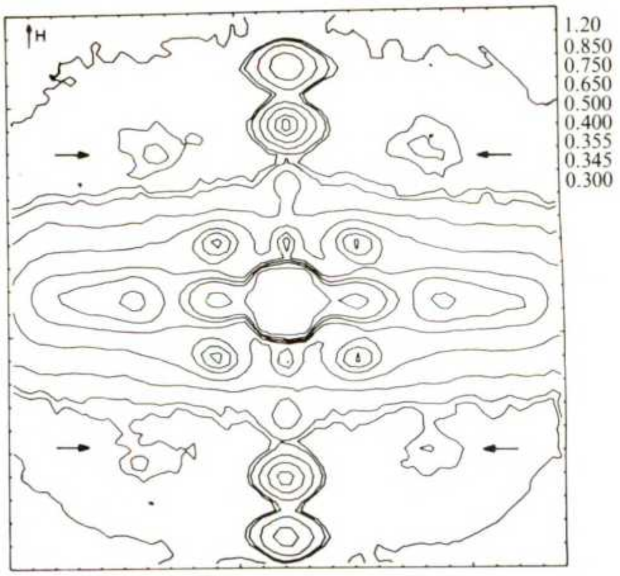


Fig. 3 Neutron diffraction pattern from fibrin (14 mg ml^{-1}) oriented in a magnetic field H of up to 18 T (arrow indicates field direction). The contour lines connect points of equal diffraction intensity. The conditions of the reaction are analogous to those described in Fig. 1 legend. The orientation was carried out in H_2O buffer which was replaced with D_2O buffer by diffusion, thereby greatly reducing the background level. The meridional reflections have a 223 ± 5 Å periodicity and the sharp equatorial peaks on either side of the centre are at 184 ± 5 Å. The arrows are drawn midway between the second and third layer lines. The central, rectangular region is due to the beam catcher. A sample in H_2O buffer gave a similar but less detailed diffraction pattern. Much slower polymerization in a field of 7 T also produced substantial orientation. The data were obtained with the two-dimensional (64 × 64-cm) detector of the low-angle scattering instrument D11[26] at the Institut Laue-Langevin, Grenoble. The specimen-to-detector distance was 2.55 m, the wavelength 5.8 Å and the spread $\Delta\lambda/\lambda$ 8% (full width, half maximum). So far, X-ray patterns from gels formed as above have only shown the meridional reflections.

The magnetic orientation demonstrated above is due to the diamagnetic anisotropy of the particles[9]. This depends strongly on the orientational order of the constituent diamagnetically anisotropic groups, which are chiefly α-helical peptide bonds and aromatic amino acid residues. Fibrin fibres orient parallel to the magnetic field and they contain a substantial proportion of α-helix[14] (~30% (refs 15, 16) by analogy with fibrinogen), which is also known to orient in the field direction[9,17]. A preliminary conclusion, which will be justified more completely elsewhere, is therefore that the α-helical regions in fibrin are aligned relatively parallel to the fibre long axis.

Clearly magnetic field orientation is a valuable technique. This is demonstrated by the success in orienting cell membranes[18,20] and two cylindrical viruses[21,22]. The experiments reported here confirm that magnetic orientation of biological samples is increasingly viable[21]. The technique could be extended to other biopolymers and to living cells. Oriented polymers have previously been produced by the application of magnetic fields[23–25], but in these the monomers formed a liquid-crystalline phase which oriented before polymerization advanced. In the formation of fibrin, the fibrinogen molecules (the precursors of monomers) interact very little, do not show

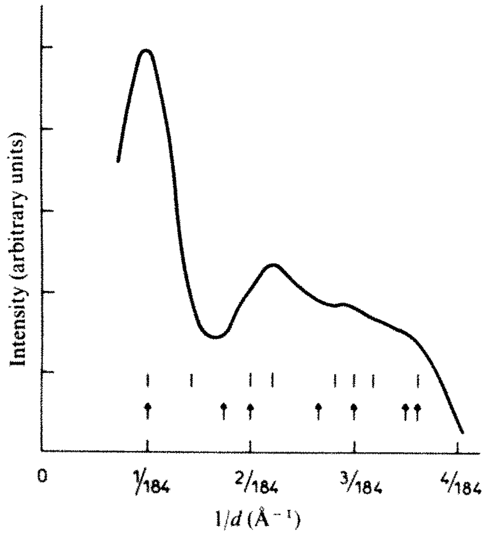


Fig. 4 The distribution of neutron-scattered intensity along the equator of Fig. 3; $1/d$ is the reciprocal of the Bragg spacing. The possible equatorial reflections from unit cells with a square (side 184 Å) and hexagonal (side $184 \times 2/3^{1/2}$ Å) base are shown by vertical lines and arrows respectively.

liquid–crystalline properties and orient poorly even in the strongest fields. These are important differences—they demonstrate that a liquid–crystalline phase or preformed orientation is not a prerequisite for the production of oriented polymers.

We thank G. Maret and C. Wilkinson for help and C. Berthet-Colominas for help with X-ray diffraction experiments. J.T. thanks the Deutsche Forschungsgemeinschaft for support. The work was partly supported by INSERM, FRA21. We are indebted to the Service National des Champs Intenses and to the Institut Laue-Langevin, both in Grenoble.

Received 28 July; accepted 24 October 1980.

1. Doolittle, R. F. *Adv. Protein Chem.* **27,** 1–109 (1973).
2. Doolittle, R. F. in *The Plasma Proteins* Vol. 2 (ed. Putnam, F. W.) 109–161 (Academic, New York, 1975).
3. Hall, C. E. & Slayter, H. S. *J. biophys. biochem. Cytol.* **5,** 11–15 (1959).
4. Fowler, W. E. & Erickson, H. P. *J. molec. Biol.* **134,** 241–249 (1979).
5. Tooney, N. M. & Cohen, C. *J. molec. Biol.* **110,** 363–385 (1977).
6. Weisel, J. W., Warren, G. S. & Cohen, C. *J. molec. Biol.* **126,** 159–183 (1978).
7. Krakow, W., Endres, G. F., Siegel, B. M. & Scheraga, H. A. *J. molec. Biol.* **71,** 95–103 (1972).
8. Ferry, J. D. *Proc. natn. Acad. Sci. U.S.A.* **38,** 566–569 (1952).
9. Worcester, D. L. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5475–5477 (1978).
10. Stryer, L., Cohen, C. & Langridge, R. *Nature* **197,** 793–794 (1963).
11. Conio, G., Dondero, G., Troglia, C., Trefiletti, V. & Patrone, E. *Biopolymers* **14,** 2363–2372 (1975).
12. Karges, H. E. & Kühn, K. *Eur. J. Biochem.* **14,** 94–97 (1970).
13. Hermans, J. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1189–1193 (1979).
14. Bailey, K., Astbury, W. T. & Rudall, K. M. *Nature* **151,** 716–717 (1943).
15. Mihalyi, E. *Biochim. biophys. Acta* **102,** 487–499 (1965).
16. Budzynski, A. Z. *Biochim. biophys. Acta* **229,** 663–671 (1971).
17. Torbet, J. & Maret, G. *Biopolymers* (submitted).
18. Geacintov, N. E., Van Nostrand, R., Pope, M. & Tinkel, J. B. *Biochim. biophys. Acta* **226,** 486–491 (1972).
19. Saibil, H. R., Chabre, M. & Worcester, D. J. *Nature* **262,** 266–270 (1976).
20. Neugebauer, D. -Ch., Blaurock, A. E. & Worcester, D. L. *FEBS Lett.* **78,** 31–35 (1977).
21. Torbet, J. & Maret, G. *J. molec. Biol.* **134,** 843–845 (1979).
22. Maret, G. *et al.* in *Nonlinear Behaviour of Molecules, Atoms and Ions in Electric, Magnetic or Electromagnetic Fields* (Elsevier, Amsterdam, 1979).
23. Cser, F. *J. Phys. théor. appl.* **40,** 459–470 (1979).
24. Perplies, E., Ringsdorf, H. & Wendorff, J. H. *Polym. Lett. Ed.* **13,** 243–246 (1975).
25. Clough, S. B., Blumstein, A. & Hsu, E. C. *Macromolecules* **9,** 123–127 (1976).
26. Ibel, K. *J. appl. Crystallogr.* **9,** 630–643 (1976).

Orientation of oxygen in oxyhaemoproteins and its implications for haem catabolism

Stanley B. Brown*, Anthony A. Chabot†, Elizabeth A. Enderby* & Anthony C. T. North†

* Department of Biochemistry and † Astbury Department of Biophysics, University of Leeds, Leeds LS2 9JT, UK

Haem is degraded to bile pigments in the catabolism of haemoproteins in mammals[1] and in the formation of photosynthetic pigments in algae[2]. The first stage of this reaction involves oxygen attack at one of the four methene-bridge carbon atoms, which is ultimately eliminated as CO(ref. 1). The four bridges are not sterically equivalent (Fig. 1) and the bilirubin in mammalian bile and algal bile pigments consists almost exclusively of the α-isomers. Little is known about the structures of the ring-cleaving enzymes responsible, although microsomal haem oxygenase, which catalyses the breakdown of haem to biliverdin in mammals[1], has very similar spectroscopic properties to myoglobin[3]. The degradation process has been simulated *in vitro* by a 'coupled oxidation' method in which the proportions of the four possible isomeric products depend on the nature of the globin moiety to which the haem is bound[4]. We report here the use of an interactive computer display system[5] to explore the relative accessibilities of the four methene bridges to a haem-bound oxygen molecule in myoglobin and in the α and β chains of haemoglobin. Our calculated interaction energies agree well with the proportions of the four isomers that are observed experimentally.

The coupled oxidation method, in which the haem compound is reacted with a reducing agent (for example, ascorbate or hydrazine) and molecular oxygen[4], seems to be a good model for biological haem cleavage as the products (biliverdin, carbon monoxide and free iron) are the same and a reducing agent (NADPH) and molecular oxygen are required *in vivo* also. Labelling experiments using ^{18}O have confirmed the similarity of the mechanisms of oxygen incorporation into bile pigment in living rats[6], algae[7], in an *in vitro* haem oxygenase system[8] and in various coupled oxidation systems[8,9].

Whereas coupled oxidation of free haem in aqueous pyridine produces an almost random mixture of the four isomers[4], coupled oxidation of haemoproteins results in selective reactions; myoglobin, like haem oxygenase, yields α isomer almost exclusively, whereas human haemoglobin yields 65% α isomer and 35% β isomer with little or no γ or δ isomer[4]. After examination of molecular models of haemoglobin[10] and myoglobin[11], Brown[12] suggested an intramolecular mechanism for the activation of molecular oxygen for haem degradation and for the determination of bridge selectivity, an essential feature of which is attack of a methene bridge carbon atom by an iron-bound oxygen molecule, considered to be liganded as in oxyhaemoglobin or oxymyoglobin. In protein-free haem, each methene bridge would be equally accessible—this would account for the random formation of biliverdin isomers. In haemoproteins, access might be restricted by the limited space in the haem pocket, and the molecular models suggested that variations in accessibility were broadly consistent with the ratios of products observed.

We attempted to evaluate this intramolecular degradation hypothesis in a more quantitative way by calculating the energies of interaction between a haem-bound oxygen molecule and the surrounding protein residues as the oxygen molecule is rotated about the Fe-O axis (Fig. 1). We used an interactive computer display system[5] in which the present molecular conformation was depicted on a cathode ray tube as it was adjusted, and which could also evaluate functions that depend on conformation, such as the potential energy due to dispersion forces between pairs of atoms: this we modelled by the expression $E = -a_{ij}r_{ij}^{-6} + b_{ij}r_{ij}^{-12}$, where r_{ij} is the distance between atoms i and j and a_{ij}, b_{ij} are constants characteristic of the types of atoms concerned[13]. We evaluated E at 5° steps of the torsion angle τ

028-0836/81/020093—03$01.00

which defines rotation of O_AO_B about the FeO_A bond, including contributions for all significant pairwise interactions that involve O_B, apart from those with Fe and with O_A, which are independent of τ. Energies for the two kinds of haemoglobin chain were evaluated in isolation because interactions between O_B and atoms in neighbouring subunits are insignificant. Hydrogen atoms were not included in our calculations as experience suggests that they have relatively little effect due to the 'softness' of their potential function and the ability of methyl groups to rotate so as to minimize steric hindrance.

For haemoglobin, the X-ray coordinates of the oxy derivative are not known; therefore we used those published for horse aquo-methaemoglobin, apart from those of the Fe atom. Horse and human haemoglobins have identical amino acid sequences in the haem pocket region except for the substitution of β-E14 Ser (horse) by Ala (human); we allowed for this substitution by omitting the hydroxyl group, although this residue is too remote for it to contribute significantly to the interaction with O_B. Similarly, for myoglobin we used the published coordinates of sperm whale metmyoglobin apart from those of the Fe atom. These globin coordinates were obtained from the Protein Data Bank[14], files 2MHB and 1MBN respectively. The coordinates of each of these structures had been refined with constraints that forced the atoms of the haem groups (apart from the pendant side chains) to be coplanar and 4-fold symmetrical. We placed the Fe atoms on the axis of 4-fold symmetry but displaced from the plane of the haem group towards the proximal histidine residue by the distances given by Phillips[15]: 0.033 nm for myoglobin, 0.007 nm for the haemoglobin α chain and 0.021 nm for the β chain. For each chain, we used Phillips' values of 0.19 nm for FeO_A, 0.14 nm for O_AO_B and 121° for the angle FeO_AO_B.

The energy values thus calculated (Fig. 2 *a–c*) indicate the relative probability of finding the oxygen molecule in any particular orientation τ about the haem normal; the higher the energy, the less likely is the oxygen to occupy that position. The most probable position would correspond to the energy minimum, but because we are interested in the likelihood of the oxygen being aligned so as to attack one of the four methene-bridge carbon atoms, our primary concern is with the relative energies in the neighbourhoods of the four corresponding values of τ.

For myoglobin (Fig. 2*a*) the energy shows a pronounced maximum at about 325° and at the γ bridge the energy is 105 kcal mol^{-1} above the minimum. This high energy is due to the presence of the distal His residue (E7) which renders the γ bridge completely inaccessible to the bound oxygen molecule. Of the other three bridge positions, α has by far the lowest energy (Table 1) as it is only 10° away from the minimum in a relatively flat part of the curve. The energies corresponding to the δ and β positions are much higher because of interactions with E11 Val and CD1 Phe respectively, suggesting that the β and δ orientations would be adopted only by a very few molecules. Thus, degradation of the α methene bridge would be very strongly preferred over degradation of the other bridges, which agrees with experiments.

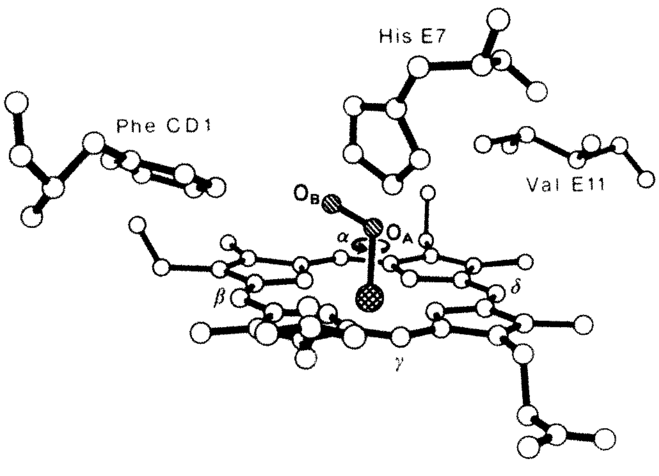


Fig. 1 Oxygen binding site in globins. The side chains shown are those primarily involved in interactions with the oxygen molecule and occur in generally similar, though not identical, positions in the three globin chains studied. The arrow indicates the torsion angle τ through which the oxygen molecule is rotated about the Fe–O_A bond. The four methene bridges are labelled α, β, γ, δ.

Table 1 Contributions made by subunits of haemoproteins to isomers formed by coupled oxidation

Haemoprotein	% Isomer formed by coupled oxidation				Energy above lowest minimum (kcal mol^{-1})			
	α	β	γ	δ	α	β	γ	δ
Sperm whale myoglobin	100	—	—	—	0.25	14.1	105	9.6
Human haemoglobin (α subunits)	55	45	—	—	0	0.6	129	10.8
Human haemoglobin (β subunits)	75	25	—	—	0	2.0	1,240	7.7

Results for α and β subunits are from ref. 160

The profile for the haemoglobin α chain (Fig. 2*b*) is grossly similar to that for myoglobin with a very high barrier due to the distal His and a relatively high energy corresponding to the δ bridge. The low energy region from $\tau = \sim 160$ to $\tau = 280°$ is much broader than in myoglobin and encompasses both the α and β positions. The actual minimum is exactly at the α bridge but the energy at the β position is only 0.6 kcal mol^{-1} higher. This suggests that both α and β bridges would be susceptible to attack, with the α bridge marginally favoured.

For the β subunits also, the γ and δ positions show a very high energy and the lowest minimum occurs precisely at the α bridge. The region between $\tau = 160$ and $\tau = 280°$ is different in profile from that for the α subunit, with an energy barrier between the α and β methene bridges. At the β bridge, the energy is 2 kcal mol^{-1} above that at the α bridge, so that more α isomer and less β isomer would be formed from haemoglobin β chains than from α chains. However, taken together, our calculations are clearly consistent with the experimental observation that the biliverdin produced by coupled oxidation of whole haemoglobin contains 65% α and 35% β isomer.

While these calculations were being carried out, Docherty[16] determined the separate contributions made by the α and β subunits to the isomer pattern obtained by coupled oxidation of haemoglobin. His results (Table 1), which showed that the β subunits produce some β isomer but not as much as do the α subunits, conformed well with our predictions.

Our observations strongly favour the intramolecular theory of haem degradation both in explaining the absence of γ and δ isomers and in accounting for the variation in proportion of the α and β isomers. Further support has been obtained from studies of the coupled oxidation of some abnormal haemoglobins[17]. For example, in haemoglobin Zürich, E7 His (the residue that blocks access to the γ bridge) is replaced by Arg, which is displaced from the haem pocket. We therefore predict that the γ methene bridge might be accessible to attack, which agrees with the experimental result that the γ isomer is present in the products.

Our energy profiles vary rapidly in the region of the β bridge so that relatively small changes in protein structure can greatly affect the amount of β isomer produced; this may explain the variations observed in the amount of β isomer in the degradation products of animal haemoglobins (S.B.B., unpublished) and plant leghaemoglobins[18].

There are several uncertainties in our calculations which cannot be readily resolved: (1) the coordinates used refer strictly to the aquo-met forms of the haemoproteins rather than the oxy forms; (2) the geometry of the oxygen molecule in the transition state of the reaction may differ from that assumed; (3) the

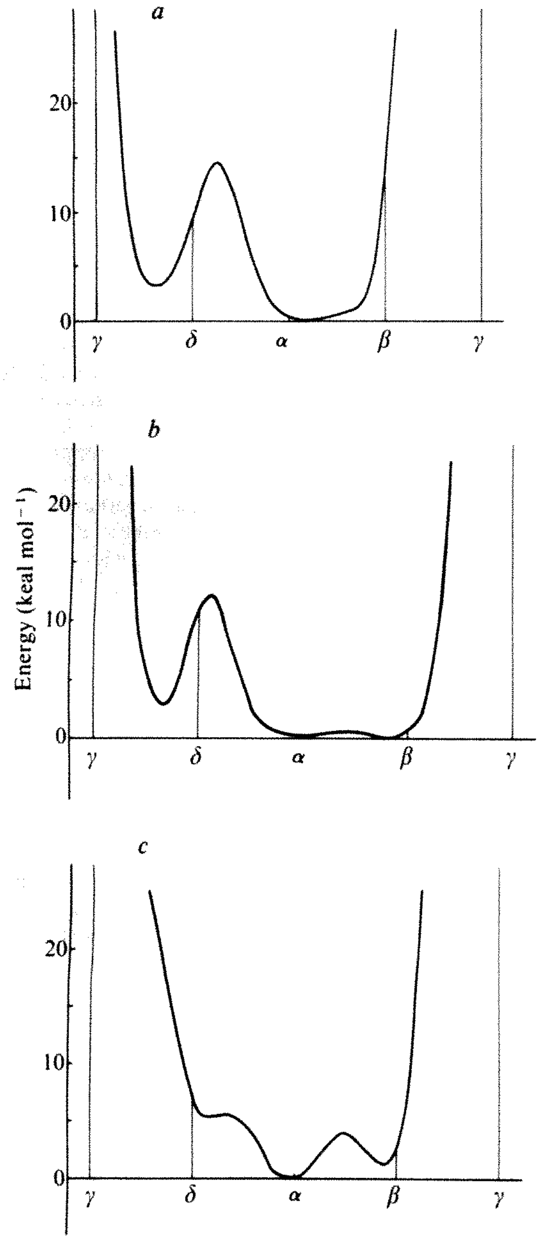


Fig. 2 Variation of interaction energy with rotational angle τ for several haemoproteins. The energy values refer to an arbitrary zero corresponding to the minimum in each case. Rotation is such that the α, β, γ and δ bridges correspond to τ values of 180, 270, 360(0°) and 90° respectively. *a*, Sperm whale myoglobin; *b*, human haemoglobin (α subunits); *c*, human haemoglobin (β subunits).

oxygen may be suitably disposed to attack a methene bridge within a range of τ values, necessitating integration of the energy functions; (4) we ignored electrostatic interactions and (5) it is now known (S.E.V. Phillips, personal communication) that the haem rings are not coplanar but rather puckered. Clearly, we would not have obtained the present level of correlation if these effects were too large, but we would be cautious about a more quantitative interpretation until data on other haemoproteins become available.

Finally, in all three haemoprotein chains studied, the minimum energy position is coincident with or close to the position of the α methene bridges. As there is no intrinsic reason that the energy minimum should correspond to any of the methene bridges, we wonder whether the situation is advantageous. It seems unlikely, although not impossible, that this feature of haemoproteins has evolved to facilitate the production of α isomers of bile pigments after degradation. Perhaps an 'open' α methene-bridge region may favour access of oxygen to the iron atom in the normal reversible interaction of haemoproteins with oxygen.

We thank the MRC for the award of a Project Grant to S.B.B. and the SRC for support to A.C.T.N. A.A.C. was supported by a Wellcome Research Studentship.

Received 25 July; accepted 10 November 1980.

1. Schmid, R. & McDonagh, A. F. *Ann. N.Y. Acad. Sci.* **244,** 533–552 (1975).
2. Troxler, R. F. & Bogorad, L. *Pl. Physiol.* **41,** 491–499 (1966).
3. Yoshida, T. & Kikuchi, G. *J. biol. Chem.* **254,** 4487–4491 (1979).
4. O'Carra, P. in *Porphyrins and Metalloporphyrins* (ed. Smith, K. M.) 123–153 (Elsevier, Amsterdam, 1975).
5. North, A. C. T., Denson, A. K., Evans, A. C., Ford, L. O. & Willoughby, T. V. in *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (ed. Srinivasan, R.) 59–72 (Pergamon, Oxford, 1980).
6. Brown, S. B. & King, R. F. G. J. *Biochem. J.* **170,** 299–311 (1978).
7. Troxler, R. F., Brown, A. S. & Brown, S. B. *J. biol. Chem.* **254,** 3411–3418 (1979).
8. King, R. F. G. J. & Brown, S. B. *Biochem. J.* **174,** 103–109 (1978).
9. Chaney, B. D. & Brown, S. B. *Biochem. Soc. Trans.* **6,** 419–421 (1978).
10. Ladner, R. C., Heidner, E. G. & Perutz, M. F. *J. molec. Biol.* **114,** 385–414 (1977).
11. Kendrew, J. C. & Watson, H. C. *Prog. Stereochem.* **4,** 299–333 (1969).
12. Brown, S. B. *Biochem. J.* **159,** 23–27 (1976).
13. Scott,, R. A. & Scheraga, H. A. *J. chem. Phys.* **45,** 2091–2101 (1966).
14. Bernstein, F. C. *et al. J. molec. Biol.* **112,** 535–542 (1977).
15. Phillips, S. E. V. *Nature* **273,** 247–248 (1978).
16. Docherty, J. C. thesis, Univ. Leeds (1980).
17. Brown, S. B. & Docherty, J. C. *Biochem. J.* **173,** 985–987 (1978).
18. Lehtovaara, P. & Perttila, U. *Biochem. J.* **176,** 359–364 (1978).

Are archaebacteria merely derived 'prokaryotes'?

Carl R. Woese & Ramesh Gupta

Department of Genetics and Development, 515 Morrill Hall,
University of Illinois, Urbana, Illinois 61801

The archaebacteria are a group of prokaryotes which seem as distinct from the true bacteria (eubacteria) as they are from eukaryotes[1–4]. The evidence on which this conclusion rests is of two types: genotypic (quantitative)—that is, comparative sequence studies, and phenotypic (qualitative)—that is, differences in various organismal characteristics. The differences between archaebacteria and true bacteria are so great, both quantitatively and qualitatively, that the two bacterial groups should be considered as representing separate primary lines of descent, each tracing directly back to the universal ancestor[1,4]. Furthermore, this ancestor itself seems not to be a prokaryote; rather it was a far simpler type of organism, one properly called a progenote[5,6]. If this is true, the discovery of archaebacteria marks a major advance in the biologist's attempts to understand the basis for the evolution of the cell.

Recently van Valen and Maiorana have challenged this interpretation[7]—they consider that the archaebacteria represent merely one branch of a general prokaryotic phylogenetic tree. Their interpretation requires the archaebacteria to have undergone more extensive evolution, both quantitatively and qualitatively, than have other prokaryotes. They also claim that archaebacteria are the forerunners of the so-called eukaryotic 'host cell', the presumed entity which housed the endosymbionts that later became eukaryotic mitochondria, chloroplasts, etc.[7]. Their interpretation is a conservative one—it seeks to retain the conventional prejudices regarding the evolution of bacteria and the origins of the eukaryotic cell (for example, prokaryotes are monophyletic; fermentative metabolism preceded autotrophic or phototrophic metabolism in prokaryotes; there are only two basic types of living system on Earth, eukaryotes and prokaryotes; the eukaryotic ancestor was a prokaryote that lost its cell wall[1,7,8–10]). A clearcut choice between the conservative and the radical interpretation of the facts cannot be made without more evidence than now exists. Our purpose in discussing the two interpretations, therefore, is

0028-0836/81/020095—02$01.00

not to eliminate one, but to stress the differences between them, thus preparing the ground for a future decision when more becomes known about the archaebacteria.

First, consider the genealogical relationship of archaebacteria to true bacteria ('prokaryotes' as the molecular biologist defines them). The issue is whether archaebacteria constitute a line of descent within the phylogenetic tree defined by the true bacteria[7], or whether the archaebacterial line branched from the common stem at a point before the ancestor of all true bacteria had evolved, and if the latter, whether the ancestor common to true bacteria and archaebacteria was itself a prokaryote or a simpler form, a progenote in our terminology[1,6]. Are prokaryotes of monophyletic or polyphyletic origin?

Those properties shared by (most) true bacteria must be largely ancestral. When defined in moderate detail many of these presumed ancestral properties are found to be absent in archaebacteria, which instead have different but equivalent properties. The conservative view is that archaebacteria have lost these ancestral 'prokaryotic' properties and have re-evolved their equivalents: for example, archaebacteria (on the basis of this conservative view) would have had to lose the typical bacterial cell wall and evolved new types of walls[3,5,11,12]; have drastically altered the characteristic sequence of 'prokaryotic' rRNA (ref. 3); changed the characteristic pattern for nucleotide modification in both rRNA and tRNA[2,4,13]; evolved ether-linked, branched chain lipids, discarding ester-linked, straight chain lipids in the process[14,15]; changed the subunit structure of RNA polymerase[16,17]; and they would have had to discard many of the antibiotic sensitivities that are typical of prokaryotes and of eukaryotic organelles, again acquiring different ones[2]. Were archaebacteria as well characterized as are the true bacteria, this list of changes in true bacterial characteristics, ostensibly wrought during archaebacterial evolution, might be enormous. If we extrapolate from the little knowledge we have of archaebacteria, it seems that true bacteria and archaebacteria differ significantly in the details of almost all molecular processes[2,4].

The conservative explanation attempts to account for the many molecular disparities between archaebacteria and true bacteria by invoking a period of 'rapid evolution' at an early stage in the line of archaebacterial descent[7]. This explanation is too plastic and ill-defined; seemingly any outcome can be rationalized. Peculiarities of archaebacterial niches alone could not seemingly account for this supposed rapid evolution. Species of eubacteria cohabit some niches with archaebacteria; for the former, there is evidence of only minor, local changes in molecular processes as a result of adaptations[18]. Examples of apparently rapid evolution among true bacteria have been encountered—the true mycoplasmas are one. (These organisms are actually highly derived offshoots of the clostridia[19].) Although they are unique, the mycoplasmas retain a great deal more genotypic resemblance to true bacteria than do the archaebacteria[19]. The eukaryotic organelles are another example[20,21]—in the case of the animal mitochondrion perhaps more drastic than might be expected from a free-living organism. However, even the highly idiosyncratic animal mitochondrion, which has the least rRNA sequence homology with eubacterial rRNAs of all organisms[5,20], shows a marked resemblance to true bacteria—for example, in the details of its respiratory chain[22], antibiotic sensitivities[23], modified nucleotides in tRNA[24], and perhaps in the secondary structure of its rRNA (H. F. Noller and C. R. Woese, unpublished).

The issue here is the extent of the differences between archaebacteria and true bacteria. Is it possible that once the cell's basic molecular processes have been established (and are well integrated into the fabric of the cell), that an organism can proceed along an evolutionary course and encounter niches that cause it to alter almost all these processes in significant ways? We consider the amount of change that can be effected in the basic molecular processes once they are established is quite limited, and if our present knowledge of archaebacteria does not exceed these limits, then our future knowledge of them certainly will. For this reason, it has been suggested that the archaebacterial and true bacterial lines of descent separated from one another during the stage at which the cell's basic molecular processes were still evolving and being refined[4-6]. Each line separately completed the evolution of the various molecular functions, as it evolved from the progenote to the prokaryotic state[1]. (One of the predictions of the radical view is that archaebacteria should possess very different genetic control mechanisms from true bacteria[2,6], in that genetic control is largely a series of refinements in basic gene expression mechanisms.)

The conservative interpretation is that the archaebacterial line of descent gave rise to the eukaryotic cell, that is, the putative host in which the endosymbionts reside[7]. It is not unreasonable that archaebacterial genes and structures are found in the chimaera that is the eukaryotic cell, like true bacterial genes and structures are[4] (C.R.W., in preparation). In fact, one of the archaebacterial ribosomal proteins shows a distinct specific sequence homology with its eukaryotic counterpart[25]. Archaebacteria do share a number of properties specifically with eukaryotes[2,4,26,27]. However, we consider that such a suggestion is unwarranted, for several reasons. Few facts are known and they do not strongly support the argument. Archaebacteria have as many properties in common with true bacteria as they do with eukaryotes, and true bacteria also share properties specifically with eukaryotes[2,28,29]; also the eukaryotic 18S rRNA seems to be no more closely related to its archaebacterial than its eubacterial counterpart[1,4]. Furthermore, archaebacteria already have a disturbing number of dissimilarities to eukaryotes[2]. Too little is known about the way in which eukaryotic cells evolved to justify this specific conclusion[7]—we do not know enough about genealogies of various eukaryotic (and bacterial) genes. We are not even certain that there was such a thing as the ancestral eukaryotic 'host cell', as it is now envisaged[7,8]. We do not know the extent to which endosymbioses are responsible for the chimaeric condition of the eukaryotic nucleus (C.R.W., in preparation). It is more useful at this stage to state merely that archaebacteria as well as true bacteria seem to be represented in the eukaryotic chimaera, and that other lines of descent, other groups, may be represented there as well (for example, the 18S rRNA)[2,4]. The ways in which the various lines are so represented, the evolutionary stages at which the associations occurred and the number of such events involved are all open questions, to be decided by comparative molecular analyses of eukaryotes, eubacteria and archaebacteria.

Received 9 October; accepted 10 November 1980.

1. Woese, C. R. & Fox, G. E. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5088–5090 (1977).
2. Woese, C. R., Magrum, L. J. & Fox, G. E. *J. molec. Evol.* **11,** 245–252 (1978).
3. Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R. & Wolfe, R. S. *Microbiol. Rev.* **43,** 260–296 (1979).
4. Fox, G. E. *et al. Science* **209,** 457–463 (1980).
5. Woese, C. R. *Symp. Soc. gen. Microbiol.* **20,** 39–54 (1970).
6. Woese, C. R. & Fox, G. E. *J. molec. Evol.* **10,** 1–6 (1977).
7. VanValen, L. M. & Maiorana, V. C. *Nature* **287,** 248–250 (1980).
8. Margulis, L. *Origin of Eucaryotic Cells* (Yale University Press, New Haven, 1970).
9. Stanier, R. Y. *Symp. Soc. gen. Microbiol.* **20,** 1–38 (1970).
10. Murray, R. G. E. in *Bergey's Manual of Determinative Bacteriology* 8th edn (eds Buchanan, R. E. & Gibbons, N. E.) (Williams & Wilkins, Baltimore, 1974).
11. Kandler, O. & König, H. *Archs Microbiol.* **118,** 141–152 (1978).
12. Kandler, O. *Naturwissenschaften* **66,** 95–107 (1979).
13. Gupta, R. & Woese, C. R. *Curr. Microbiol.* (in the press).
14. Langworthy, T. A. *Biochim. biophys. Acta* **487,** 37–47 (1977).
15. Tornabene, T. G. & Langworthy, T. A. *Science* **203,** 51–53 (1979).
16. Zillig, W., Stetter, K. O. & Janekovic, P. *Eur. J. Biochem.* **96,** 597–604 (1979).
17. Zillig, W., Stetter, K. O. & Tobein, M. *Eur. J. Biochem.* **91,** 193–199 (1978).
18. Langworthy, T. A., Mayberry, W. R. & Smith, P. F. *Biochim. biophys. Acta* **431,** 550–569 (1976).
19. Woese, C. R., Maniloff, J. & Zablen, L. B. *Proc. natn. Acad. Sci. U.S.A.* **77,** 494–498 (1980).
20. Eperon, I. C., Anderson, S. & Nierlich, D. P. *Nature* **286,** 460–467 (1980).
21. Bonen, L. & Doolittle, W. F. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2310–2314 (1975).
22. John, P. & Whatley, F. R. *Nature* **254,** 495–498 (1975).
23. Flavell, R. *Biochem. Genet.* **6,** 275–291 (1972).
24. Heckman, J. E. *et al. Cell* **13,** 83–95 (1978).
25. Matheson, A. T., Moller, W., Amons, R. & Yaguchi, M. in *Ribosomes: Structure, Function, and Genetics* (eds Chamblis, G. *et al.*) 297–315 (University Park Press, Baltimore, 1980).
26. Kessel, M. & Klink, F. *Nature* **287,** 250–251 (1980).
27. Kwok, Y. & Wong, J. T-F. *Can. J. Biochem.* **58,** 213–218 (1980).
28. Steitz, J. A. *Nature* **273,** 10 (1978).
29. Woese, C. R. & Fox, G. E. *Nature* **273,** 101 (1978).

MATTERS ARISING

Pyroclastic sulphur

FRANCIS *ET AL.*[1] recently commented on the apparent rarity of pyroclastic sulphur in terrestrial eruptions. They described sulphur-rich ejecta from Poás volcano, Costa Rica, and mentioned the presence of myriads of small, sulphur spheres. Similar spheres have been reported in tephras erupted from Ruapehu[2–4], New Zealand, and may well be common at many volcanoes where an active vent is occupied by a crater lake.

The Ruapehu spheres range from 50 μm to 1 mm, but most are 125–750 μm in diameter. They have a lustrous, granular surface, and comprise solid, black, orthorhombic sulphur with disseminated silicate, and iron sulphide dust, a few small vacuoles and rare, yellow sulphur crystals. Some are globular, some are fused double-spheres, but most are almost perfectly round 'rough[5] spheres'. They were ejected in abundance by phreatomagmatic eruptions in 1971, and one fine ash collected 15 km downwind comprised 12% by volume of spheres[2]. Sulphur also occurs in Ruapehu tephras as uncommon, angular, yellow fragments, and sulphur drip on ejected blocks. Rafts of spongy sulphur float on the lake in quiet or eruptive periods.

Vent morphology and activity of Ruapehu[6] and Poás appear similar, and a subaqueous liquid sulphur lake has been postulated for both volcanoes[7,8]. I agree with Francis *et al.*[1] that this is unlikely. Much consolidated, volcaniclastic/chemogenic sediment has been erupted from Ruapehu, and in some siltstones ejected in 1971, sulphur occurs as globules in cavities, and veins crossing the bedding. This I believe is evidence that sulphur is 'sweated' out of sulphurous sediment, whence it may enter superheated steam fumaroles and be discharged violently into muddy lake water to form black droplets. The lack of coarse, clastic or scoriaceous sulphur suggests that there were no large accumulations of liquid sulphur present when the lake bed was fragmented during phreatomagmatic eruptions of Ruapehu.

C. P. WOOD

New Zealand Geological Survey,
PO Box 30368,
Lower Hutt,
New Zealand

1. Francis, P. W., Thorpe, R. S., Brown, G. C. & Glasscock, J. *Nature* **283**, 754–756 (1980).
2. Wood, C. P. *N.Z. Volcan. Rec.* **1**, 22–24 (1973).
3. Wood, C. P. *N.Z. D.S.I.R. Bull.* **218**, 103–108 (1977).
4. Wood, C. P. *N.Z. D.S.I.R. Bull.* **224**, 35–39 (1978).
5. Walker, G. P. L., Wilson, L. & Bowell, E. L. G. *Geophys. J. R. astr. Soc.* **22**, 377–383 (1971).
6. Cole, J. W. & Nairn, I. A. *Catalogue of the Active Volcanoes and Solfatara Fields of New Zealand* (IAVCEI, Rome, 1975).
7. Giggenbach, W. *N.Z. J. Sci.* **17**, 33–45 (1974).
8. Bennett, F. D. & Raccichini, S. M. *Nature* **271**, 342–344 (1978).

FRANCIS REPLIES—We were not aware of the existence of small sulphur spherules in the material ejected by the Ruapehu eruption of 1971, and are grateful to Wood for bringing it to our attention. We agree that such spheres may well be common in active volcanoes with crater lakes, and would welcome any further observations of these and related phenomena.

P. W. FRANCIS

Department of Earth Sciences,
The Open University,
Walton Hall,
Milton Keynes MK7 6AA, UK

Recombination intermediates

IN the summary of their elegant demonstration of adenovirus recombination intermediates, Wolgemuth and Hsu[1] state: "To the best of our knowledge, this is the first visualization of *in vivo* recombination intermediates of discrete DNA molecules isolated from eukaryotic cells". They appear to have overlooked the paper published last year by Bell and Byers[2] which presented an equally clear demonstration of Holliday-type recombination intermediates involving the 2-μm plasmids of *Saccharomyces cerevisiae.* Although this earlier case might be held to be atypical of recombination in general, in that it involved inverted repeats within the same molecule rather than homologous regions of different molecules, it was, in another respect, more relevant to meiotic recombination, because the Holliday structures were found specifically in cells at the prophase stage of meiosis.

J. R. S. FINCHAM

Department of Genetics,
University of Edinburgh,
Edinburgh EH9 3JN, UK

1. Wolgemuth, D. J. & Hsu, M.-T. *Nature* **287**, 168–171 (1980).
2. Bell, L. & Byers, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3445–3449 (1979).

WOLGEMUTH AND HSU REPLY—As stated in the text of our paper, our study does indeed represent the first example, to the best of our knowledge, of visualization of *in vivo* recombination intermediates in *animal* cells and we are sorry that the word eukaryotic in the summary was misleading. We would like to emphasize Fincham's point that our observations are more typical of recombination in that they are inter- rather than intramolecular. Our failure to cite the paper by Beck and Byers in our discussion was totally inadvertent and we had contacted them previously in this regard.

DEBRA J. WOLGEMUTH
MING-TA HSU

The Rockefeller University,
1230 York Avenue,
New York, New York 10021

Common ancestry for birds and crocodiles?

A PSEUDOSUCHIAN origin for birds and crocodiles, independent of dinosaurs, has been proposed by Whetstone and Martin[1]. They described two 'derived' characters of the otic region of the skull, a fenestra pseudorotunda and periotic sinuses, which were claimed to be unique to crocodiles and birds and thus evidence for their common ancestry. It is well known that crocodiles and birds possess a similar otic morphology[2,3]. Any hypothesis of their common ancestry independent of dinosaurs therefore depends critically on the assertion that dinosaurs do not have similar shared features, or that if they do, the features are not homologous. In this respect, we think that Whetstone and Martin's evidence is equivocal and should be treated with caution.

Whetstone and Martin note that in primitive reptiles the perilymphatic duct exits from the cranium through the vagus foramen (fissura metotica), accompanied by certain cranial nerves. In crocodiles and birds, however, the perilymphatic duct has a separate exit, and terminates in a small membrane that closes the fenestra rotunda, or, more correctly, the fenestra pseudorotunda[2]. The fenestra pseudorotunda lies very close to the fenestra ovalis.

Based on their interpretations of casts of the inner ear structure, Whetstone and Martin concluded that dinosaurs did not have a fenestra pseudorotunda. Casts of four of the seven genera studied were illustrated: the sauropod *Brachiosaurus*, the hadrosaur *Lophorothon*, the ankylosaur *Ankylosaurus* and the theropod *Allosaurus* (see ref. 1, Fig. 2 *a–d*, respectively). We question how the authors can be certain that the single trunk identified in *Lophorothon* and *Ankylosaurus* does in fact represent both the perilymphatic duct and certain cranial nerves. Why should it not represent the perilymphatic duct alone? In *Lophorothon*, for example, could not the first trunk (labelled perilymphatic duct plus nerves X–XI) represent the perilymphatic duct exiting alone, and

0028–0836/81/020097—02$01.00

ending in the fenestra pseudorotunda? The next trunk could then represent nerves X and XI and the third trunk nerve XII. The point here is that Whetstone and Martin's interpretation of a complex morphological structure from casts is questionable, and alternative explanations are possible.

Notwithstanding these questions of interpretation, which simply reflect the shortcomings of attempts to reconstruct the soft anatomy of fossil organisms, it is interesting that other workers have reported a fenestra pseudorotunda in the theropods *Dromaesaurus*[4] and *Gallimimus*[5], and in the pachycephalosaurians *Prenocephale* and *Homalocephale*[6]. In *Dromaeosaurus*[4] remnants of the stapes confirm the interpretation of the fenestra ovalis, and the adjacent aperture, identified as the fenestra (pseudo) rotunda (though identified as the IX foramen by Whetstone and Martin), bears a relationship with the fenestra ovalis which is strikingly similar to that illustrated by Whetstone and Martin for *Hesperornis* and *Alligator* (compare ref. 4, Fig. 7*a*, and ref. 1, Figs. 3*a*,*b*).

Whetstone and Martin marshalled evidence from embryology to support their case that the 'round window' (fenestra pseudorotunda) of birds and crocodiles is a homologous feature unique to them. Although other living vertebrates (mammals and lizards) also have round windows, they argued that the structure had a similar embryology in birds and crocodiles, and that this differed from that of other vertebrates. However, if some dinosaurs do indeed have a fenestra pseudorotunda, then the embryological evidence in support of its homology in birds and crocodiles is less compelling. The similar embryology of the otic regions of crocodiles and birds may only reflect their archosaurian ancestry.

The other character described by Whetstone and Martin, periotic sinuses associated with bones of the middle ear cavity, also may not be unique to crocodiles and birds. Depressions have been described in the lateral wall of the braincase of *Saurornithoides*[7] that seem to be associated with the middle ear.

Thus, we find that the evidence presented by Whetstone and Martin to support an hypothesis of common ancestry of crocodiles and birds independent of dinosaurs is equivocal. It is sobering to recount Baird's[8] summary statement in 1970: 'Parallelism and convergence in the tympanic regions of living reptiles are common, and only a few reptilian ears are yet known in detail.' Thus, even if Whetstone and Martin are correct in their assertion that a fenestra pseudorotunda and periotic sinuses are unique to birds and crocodiles, it remains to be determined whether this is due to common ancestry or convergence. In view of the extensive evidence compiled by Ostrom[9-14] for common ancestry of birds and dinosaurs, the otic features mentioned by Whetstone and Martin may best be ascribed to homoplasy.

C. McGowan

Department of Vertebrate
Palaeontology,
Royal Ontario Museum,
Toronto, Ontario,
Canada MS5 2C6 and
Department of Zoology,
University of Toronto,
Toronto, Ontario,
Canada MS5 1A1

A. J. Baker

Department of Ornithology,
Royal Ontario Museum,
Toronto, Ontario,
Canada MS5 2C6
and Department of Zoology,
University of Toronto,
Toronto, Ontario,
Canada MS5 1A1

1. Whetstone, K. N. & Martin, L. D. *Nature* **279,** 234–236 (1979).
2. de Beer, G. R. *The Development of the Vertebrate Skull* (Oxford University Press, 1937).
3. Romer, A. S. & Parsons, T. S. *The Vertebrate Body* (Saunders, Philadelphia, 1977).
4. Colbert, E. H. & Russell, D. A. *Am. Mus. Nov.* No. 2380, 1–49 (1969).
5. Osmólska, H., Roniewicz, E. & Barsbold, R. *Palaeont. pol.* **27,** 103–143 (1972).
6. Maryanská, T. & Osmólska, H. *Palaeont. pol.* **30,** 46–102 (1974).
7. Barsbold, R. *Palaeont. pol.* **30,** 5–22 (1974).
8. Baird, I. L. in *Biology of the Reptilia: Morphology B.* Vol. 2 (eds Gans, C. & Parsons, T. S.) 193–275 (Academic, London, 1970).
9. Ostrom, J. H. *Nature,* **242,** 136 (1973).
10. Ostrom, J. H. *Q. Rev. Biol.* **49,** 27–47 (1974).
11. Ostrom, J. H. *A. Rev. Earth planet. Sci.* **3,** 55–57 (1975).
12. Ostrom, J. H. *Proc. Centre natn. Rech. Sci.* **218,** 519–532 (1975).
13. Ostrom, J. H. *Smithsonian Contr. Paleobiol.* **27,** 1–21 (1976).
14. Ostrom, J. H. *Biol. J. Linn. Soc., Lond.* **8,** 91–182 (1976).

WHETSTONE AND MARTIN REPLY—The fundamental difference between the otic capsules of *Sphenodon* and those of a bird or crocodile is in the course of the perilymphatic duct. In *Sphenodon*, it turns medially, enters the endocranium and passes along a groove before being exposed to pharyngeal tissues[1]. In birds and crocodiles, it turns laterally and is exposed to the tympanic cavity. Either condition requires a correlated suite of osteological structures. We have used these structures in attempting to determine the course of the duct in dinosaurs[2]. Hadrosaurs and *Ankylosaurus*, for example, are like *Sphenodon* but unlike birds or crocodiles, in having an open 'vestibule' to the endocranium, a groove between the vestibule and vagus foramen, and no confluence between the external cranial openings. Whether or not the duct left the cranium alone (as suggested by McGowan and Baker) is irrelevant to the problem at hand. Either interpretation presumes a primitive (medial) course for the duct, along the braincase wall.

By using a bird or crocodile as a model for fossil archosaurs, some workers have restores a fenestra pseudorotunda in dinosaurs. The foramen indentified[3] as the 'round window' in *Gallimimus* is not on the otic capsule, but is far up on the paroccipt. In the type of *Dromaeosaurus*, the right otic capsule has been destroyed, along with the footplate of the stapes. The resulting hole and the remnants of a nerve canal were labelled 'Foramen Rotundum' in the original restoration[4]. The ear regions of the cited pachycephalosaurs are indeterminate. The best candidate for the presence of this structure is the ornithischian, *Hypsilophodon*. Were a 'window' present in dinosaurs, it would not lessen the embryological evidence for the homology of the fenestra in birds and crocodiles, but might indicate that the homology evolved in a more remote ancestor.

Ostrom's evidence for a theropod origin for birds has been well summarized[5]. The evidence for a 'sister-group' relationship to crocodilians, first suggested by Walker[6], is still being developed by Whetstone, but the following derived features seem to be homologous: a fenestra pseudorotunda; a pneumatic quadrate; a foramen aerosum in the lower jaw; periotic pneumatic cavities in the dorsal, central and rostral positions: a quadrate cotylus at the anterior base of the paroccipt; a bipartite quadrate articulation with dermal and endochondral bones—anteriorly with the prootic, squamosal and laterosphenoid, posteriorly with the prootic and otoccipital; a squamosal shelf over the ear region; anterio-medial origin of the temporal musculature; two pneumatic cavities surrounding the cerebral carotid; unserrated teeth with a constricted neck[7]; bony tooth roots with an enclosed, oval resorption pit[7]. None of these features is known in theropod dinosaurs. The 'lateral depression' of *Saurornithoides* has no similarity to the periotic sinuses.

On the basis of present evidence, we feel that an argument of close relationship to crocodilians is the most parsimonious hypothesis available for the ancestry of birds.

K. N. Whetstone
L. D. Martin

Department of Systematics
and Ecology,
Museum of Natural History,
University of Kansas,
Lawrence, Kansas 66045

1. Baird, I. L. in *Biology of the Reptilia: Morphology B* Vol. 2 (eds Gans, C. & Parsons, T. S.) 193–275 (Academic, London, 1970).
2. Whetstone, K. N. & Martin, L. D. *Nature* **279,** 234–236 (1979).
3. Osmólska, H., Roniewicz, E. & Barshold, R. *Palaeont. pol.* **27,** 103–143 (1972).
4. Colbert, E. H. & Russell, D. A. *Am. Mus. Nov.* No. 2380, 1–49 (1969).
5. Ostrom, J. H. *Nature* **242,** 136 (1973).
6. Walker, A. D. *Nature* **237,** 257–263 (1972).
7. Martin, L. D., Stewart, J. D. & Whetstone, K. N. *Auk* **97,** 86–93 (1980).

BOOK REVIEWS

History and non-history

Dorothy Hodgkin

Sage: A Life of J. D. Bernal. By Maurice Goldsmith. Pp.288. (Hutchinson: 1980.) £8.95.

Popperfoto

J. D. Bernal at the inauguration of a new building of the Institute for Structure Research in Germany, March 1962. Dorothy Hodgkin is in the background.

WHEN he spoke in memory of J. D. Bernal in January, 1972, Lord Snow said "I strongly advise that his biography should not be written until the end of the coming decade . . . there is so much in this rich life to learn about and to interpret". The decade is coming to an end, a large amount of material connected with Bernal's scientific and other work and some personal papers have been sorted out, and a catalogue — nearly ready — is being prepared in the archives of Cambridge University library. So, on Lord Snow's prescription, a serious book on Bernal could now be written; and, perhaps, the sooner the better, while there are still many living who knew him.

Maurice Goldsmith's book is not the book that is needed. It is not even a good forerunner of that book. It is too confused and inaccurate. Goldsmith does not understand Bernal well enough as a person, what he cared for and particularly his scientific work and achievement. This Bernal himself felt, when, semi-paralysed, he was visited by Goldsmith and learned that he wanted to write his biography. Bernal asked his wife to discourage Goldsmith from writing, which she did as far as she could, arranging for an alternative biographical study to be made and refusing Goldsmith access to the Cambridge archives. Knowing something of the state of the archives in the past few years I have every sympathy with her. Knowledge of its contents would not have removed the major defects of this book.

The book sets out to cover Bernal's life and work in 14 chapters, from his childhood in Ireland, where he was born in 1901, to his death in London in 1971. The early part is largely taken from autobiographical sketches written by Bernal, much of the rest from conversations with friends and colleagues, from newspaper reports of the time and some from Bernal's own writings. So there are many stories which bring back vivid memories of him, of his early passion for science and experiments at home, at school and Cambridge, of Sir William Bragg and the Royal Institution, and the beginning of the X-ray analysis of crystals, of Max Perutz coming to work with him at Cambridge, of his war work on air-raid precautions and with Lord Mountbatten. It was at Cambridge that Bernal acquired the nickname Sage which many of his early friends used and which gives the book its title. He read widely and to his undergraduate friends seemed extraordinarily learned. Goldsmith himself knew him best at a later stage of his life, through their common interest in the use of science, science policy, the social function of science and the science of the sciences; these are subjects dealt with at some length. There are also chapters on Bernal's relations with the Soviet Union, on the development of international organizations for science and for peace after the War, on the political controversies that troubled both his early and later career, and on his work on science in history.

Unfortunately the text is scattered with inaccuracies — in most cases easily avoidable — confusions and misconceptions. For example, it is recorded that Desmond Bernal had one brother and two sisters; he had, in fact, two brothers, Kevin and Godfrey, who are still alive; one of his two sisters died in infancy. Later the text runs: "In June 1945, Bernal was in Moscow for the 220th anniversary of the Academy of Sciences. His friend Frédéric Joliot-Curie was also there . . . together with Joliot, Bernal began discussions with a few Soviet scientists on a proposed new international body". Actually Bernal was prevented from going to the Moscow celebrations by the British government — later it is said, quite correctly, that the visit he made to Moscow in 1949 was the first since 1934. Then, about the long paper by Bernal and Fankuchen on tobacco mosaic virus: "It was most fortunate that Fankuchen finally wrote up that important paper, because Bernal would never have done so". This is invention. Anyone who knew them both would know only Bernal could have written the major part of that paper. Fankuchen was an excellent experimentalist, but his role at the end was to add a few exact figures and get the paper accepted for publication.

From childhood Bernal was passionately interested in science, in carrying out experiments himself, thinking about his own finds and those of others. As the dustcover says, "although primarily a crystallographer, he produced and provoked ideas over almost the entire range of scientific enquiries". Yet only one chapter of the 14 is devoted to the main body of his scientific work and this is extraordinarily confused and full of doubtful comments. What was really involved in his first crystallographic measurements on the sterols, and how these were linked with the work on oestrin and affected the general stream of chemical research, is never clear. Nor is the

---

Dorothy Hodgkin is Emeritus Professor at the University of Oxford, and a Fellow of Somerville College and Wolfson College, Oxford.

significance of his discovery of the conditions under which proteins and viruses could give detailed X-ray diffraction effects, work which opened the way to the elucidation of their structures and to much more, the solution of the structure of DNA and the understanding of molecular biology as we know it now. I read again the review Bernal wrote in 1968 of Watson's book, *The Double Helix*, and did not find it embarrassing as Goldsmith suggests. Bernal says specifically: "My name does not appear, and rightly, in the double helix story". That he should feel, all the same, that he had so many clues that he should have been able to see the solution himself was only natural — Pauling felt just the same on reading the book; and Bernal had an added reason in that he had himself taken X-ray photographs of nucleic acids as long ago as 1937. Incidentally, the remark that Furberg established the structure of DNA is so wrong it must be a misprint.

Other Goldsmith comments: the implication of the sentence "his own major, purely scientific contributions were made before 1939" that Bernal's work was either not his own or not major — is surely wrong. After the war Bernal set up a number of very good research groups at Birkbeck, which he left much to themselves — he strongly believed, from his early work with Sir William Bragg, that independence was a good thing — he himself working mainly on the liquid state; it was from this research that his 1959 paper on the structure of liquids emerged, arguably the most brilliant he ever wrote. Another serious misunderstanding: "he had a hurting yet non-deliberate blindness about key work done by members of his staff". Characteristically, Bernal was the most generous of colleagues, generous with ideas, help and credit (as indeed is said elsewhere). Any blindness of the kind mentioned was extremely rare with him — as with children in many large families, his scientific children sometimes felt they had less than they would have liked of their parents' attention. To most he seemed kind beyond their deserts.

Many of the failings of Goldsmith's treatment of his scientific work apply to other sections; his judgement seems to me often very wrong. He relies too much on casual, sometimes malicious, gossip. John Desmond Bernal's life and work deserves a very much more serious study and appraisal. □

More *Drosophila*

Peter Lawrence

The Genetics and Biology of Drosophila, Vol.2d. Edited by M. Ashburner and T.R.F. Wright. Pp.702. (Academic: 1980.) £54, $124.50.

THIS volume is the last in the series dealing with the genetics and biology of the individual fruit fly. Volume 3, which is in preparation, will take us away into the biology of populations, taxonomy, ecology and evolution.

I have already taken the opportunity this multi-authored work has given me to malign multi-authored works in general (*Nature* **283**, 116; 1980). One criticism of the earlier volumes was the way topics are jumbled up, which makes it necessary for the specialist to buy several volumes when he may only want to read a chapter or two in each. Another was that there appeared to be a lack of discipline which led to chapters being over-long and containing evanescent material, making the individual volumes unnecessarily bulky and expensive. Although I think Vol.2d is a legitimate target for the same attacks, it contains several good chapters which cover related topics and is therefore easier to defend.

This volume covers lampbrush chromosomes, sex determination, dosage compensation, oogenesis and spermatogenesis. There are chapters on the structure and development of the epidermis, nervous system and tracheal system and three chapters which appear eccentric in this company — on neuromuscular physiology, fly pigments, and even one on viruses.

Most chapters are strong, concentrating in the main on useful facts rather than idle discussion. There are many good things, for instance tables of enzyme activities in flies bearing different X-autosome ratios, pictures and drawings of the central nervous system and new illustrations of the larval cuticular sensilla, tables referring the reader to papers on the structure of the imaginal discs, discussion of the types of cell junctions found in epithelia, drawings of the ramifications of the tracheal system, and a very useful compendium of the effects of eye colour mutants. Examples of idle discussion of the abstruse are much of the chapters on sex determination and viruses, and the sections on the rate of oogenesis in umpteen species of *Drosophila* and on cell lineage in ommatidial development. This volume, like bits of its predecessors, will be a valuable resource. Editing huge books like this is no delight, so please raise your glasses to Ashburner and Wright! □

Peter Lawrence is on the staff of the MRC Laboratory of Molecular Biology, Cambridge, UK.

Gut reactions

G.J. Dockray

Gastrointestinal Hormones. Comprehensive Endocrinology. Edited by George B. Jerzy Glass. Pp.1008. (Raven: 1980.) $85, $115.60 outside the USA.

TEN years ago secretin, gastrin and cholecystokinin were the only hormonal peptides to have been isolated from the gut and characterized. Today we know the structure of over a dozen distinct biologically active gut peptides. The function of most of these is still uncertain. However, it seems likely that while some function as true hormones, others may act locally by diffusion to nearby cells (paracrine effects), and still others are secreted by nerves and have neuro-regulatory roles. The editor of this book has rightly taken a broad view of gastrointestinal hormones, and has included chapters on both the classical hormones (secretin, gastrin, cholecystokinin) and a wide range of other active peptides.

Gut endocrinologists have not been reticent about reviewing the rapid advances in their field, and at least five other books (generally conference proceedings) devoted to the gut endocrine system have appeared in recent years. The value of the present book lies in its comprehensive approach to the subject, and in this it lives up to the title of the series.

There are 42 chapters contributed by workers from 33 different laboratories, all active in this area. The division of the chapters into sections has been rather arbitrary and the greater part of the book is essentially a systematic coverage of the localization, isolation, characterization, biological properties, receptor binding, physiology and measurement of gut peptides. A final series of chapters deals with those "candidate" hormones which have not been isolated but whose existence is inferred from physiological studies, and also with new peptides which are in the process of being isolated from hog intestine by Viktor Mutt and his collaborators.

In the not too distant future it seems probable that we will know of 20 or 30 chemically defined and biologically distinct peptides which may control digestion. This field is obviously moving quickly, and it is therefore welcome to see that the publishers have managed to produce the book quite rapidly (some references are barely a year old). Although it is inevitable that much of the book will soon be out of date, it is still essential reading for anyone who wants to keep up with advances in this area. It will also provide an invaluable source for understanding developments not just in gut endocrinology, but also in the study of peptides in central and peripheral nerves. □

G.J. Dockray is Reader in Physiology at the University of Liverpool.

Diagenesis by numbers

R.G.C. Bathurst

Early Diagenesis: A Theoretical Approach. By Robert A. Berner. Pp.241. (Princeton University Press: 1980.) Hardback $25, £13.70; paperback $9.50, £5.25.

THIS most readable and helpful book should be welcomed by all who wish to understand chemical processes in diagenesis, be they students, teachers or research workers. Berner's approach is rigorously logical: he selects various problems in early diagenesis, shows how these may be clarified and simplified, and how to pose sensible questions about them, and, finally, indicates how to derive models that are both testable quantitatively and valuable as guides to prediction. His mode of thought combines a rigorous quantitative approach (calculus at the level of simple integration) with a careful appreciation of such seemingly unquantifiable processes as bioturbation and bacterial decay. Above all, he is concerned with the rates at which different processes work. In this book he deals only with diagenesis in the first 100m or so of burial, thus avoiding complicating factors such as higher temperatures, uplift, weathering and metamorphism.

In Part I there are five chapters devoted to theory. After a short introduction, he presents a quantitative approach to basic concepts such as rates of burial, steady-state diagenesis and diffusive and advective fluxes. Next he examines in detail the quantitative expression of compaction, rate of deposition, rate of flow of water across a sediment–water interface, molecular diffusion and the application of Fick's laws, and how to accommodate bioturbation into a quantitative model. The following two chapters cover various chemical processes that are the driving forces for diagenetic change. These include adsorption, ion exchange, radioactive decay, microbial metabolic reactions, the energetics of precipitation, nucleation, crystal growth, transport control and dissolution, and the general matter of diagenetic redistribution, with a section on causative factors in authigenic processes.

Part II is about applications and in it Berner discusses and enlarges upon work that has been carried out in real situations. He starts with a consideration of marine sediments of the continental margins; topics include diagenesis within the zone of bioturbation, relying on work done in Long Island Sound, touching upon such matters as irrigation and molecular diffusion of ammonia into burrows constantly flushed with oxygenated water, bacterial sulphate reduction, ammonia formation, phosphate diagenesis and methane formation. He then comes to pelagic (deep-sea) sediments, dealing with carbonate dissolution and the lysocline and compensation depth, burial flux and compaction, the dissolution of opaline silica, the diagenesis of suboxic organic matter and of radioisotopes, and volcanic–seawater reactions. The final chapter embraces non-marine sediments (fresh to brackish water) and treats salinity fluctuations, iron and manganese diagenesis and hypersaline sediments. There is a 13-page bibliography and an efficient index. The chapters each have running headlines and these, combined with clear paragraph headings, make it easy for the reader to find his way around.

Berner's writing is superbly clear and readers should find the book both interesting and stimulating. It excels as an exercise in controlled scientific imagination and sets a high standard for future studies of recent diagenesis. □

R.G.C. Bathurst is Professor of Geology at the University of Liverpool and of Petroleum Geology (II) at the University of Bergen.

Attractions of magnetic interactions

B. R. Coles

Interactions in Magnetically Ordered Solids. By K. P. Sinha and N. Kumar. Pp.180. (Oxford University Press: 1980.) £12.50, $36.50.

THIS short book covers a lot of ground, although the authors explain at the outset that they have restricted themselves to localized moments in insulators and semiconductors. They reveal some ambivalence, however, towards conduction electrons, recognizing their importance as mediators of magnetic interactions (albeit in a rather concentrated treatment in which the sudden appearance of the word "impurity" may confuse many readers) but rather oddly entitling their last chapter "Conduction electron magnon interaction in ferromagnetic *insulators*" — a contradiction in terms not clarified by a brief final paragraph on spin-disorder resistivity which introduces de Gennes and Friedel's treatment of magnetic metals.

The problem of over-compression of the subject matter runs right through the book, perhaps the most glaring example being a section of one and a half pages of text and three figures which is introduced as "a brief survey of the thermodynamic properties of magnetic insulators".

As a result, it is difficult to see for whom the book is intended. Any given topic is dealt with so briefly as to constitute little more than a signpost to an area of interest, any real understanding of which by a student will require extensive reading elsewhere. Furthermore, the guidance to such reading is often very limited. For example, no reference to the book of Marshall and Lovesey (from the same publishers) is given in the short chapter on neutron–magnon interactions, which devotes two pages to magnon dispersion relations and nothing (in spite of the section introduction's claim) to magnon lifetimes. Perhaps the more theoretically inclined graduate student will use this book to pick up hints about arenas in which he can flex his mathematical muscles without having to acquire much feeling for the nature of the materials in which the fascinating phenomena of magnetism manifest themselves. □

B. R. Coles is Professor of Solid State Physics at Imperial College, University of London.

Art of catalysis

D.E.W. Vaughan

Heterogeneous Catalysis in Practice. By C.N. Satterfield. Pp.416. (McGraw-Hill: 1980.) $36.30, £16.15.

IN this book Professor Satterfield has squeezed into less than 400 pages the essential characteristics of catalysis and many of the major industrial catalytic processes. He has not attempted to be either up to date (there are few references later than 1976), or comprehensive in his coverage, but instead presents overviews that give the reader an appreciation for the processes and the associated practical problems ever present in "real" systems. The "art" side of catalysis is given a fairly complete coverage. Throughout, the references are secondary rather than primary; for example, few patents are cited. Overall, the treatment is directed toward the chemical engineer rather than the chemist.

The first half of the book deals with the necessary physical–inorganic chemistry of catalysis and catalysts, with chapters on sorption, kinetics, catalyst preparation and characterization. These are brief, and often only define the terms, nomenclature and methods of catalytic chemistry. The reader is supplied with appropriate references if a deeper understanding is desired, both in the form of chapter bibliographies and a general literature overview at the end of the book. Throughout, the emphasis is on practicality, real problems and working systems. The chapter on catalyst preparation and manufacture (avoided in most books) is a reminder of the multitudinous problems that can be built into a catalyst before measurements or tests are made on it — a sobering reminder that a knowledge of the preparation details is often essential to a complete interpretation of catalyst evaluation data.

The second part of the book surveys areas of industrially important catalytic chemistry, with chapters on supported metal catalysts, acid and zeolite catalysts, catalytic oxidation, processing of petroleum hydrocarbons, and synthesis gas and associated processes. These chapters trace the development of process catalysts up to present practice, focussing on details of catalyst composition and process conditions. The chapter on synthesis gas is particularly timely in view of the renewed interest in CO and H_2 as sources of synthetic chemicals, and includes discussions of steam reforming, Fischer-Tropsch synthesis, water-gas shift reactions, methanation, methanol synthesis and ammonia synthesis. Details are generally accurate.

The final chapter on experimental methods covers the engineering details of reactor design, heat and mass transfer, and various types of reactors useful in the laboratory. This is useful to the chemist in that it is reminder of many important non-catalyst effects on experimental results.

The book is eminently readable, in the sense that the author gets to the point in clear prose, and moves on to the next subject. In addition to use in courses in catalysis, the book will provide a valuable review of the subject for the experienced worker in the field, in that its broad coverage of pratical catalyst systems makes it a good initial reference for catalysts and process conditions.

D.E.W. Vaughan is a Research Associate with the Exxon Research and Engineering Company, working in the areas of sorbent and catalytic materials.

Interactions of binary stars

R. J. Tayler

Close Binary Stars: Observation and Interpretation. Edited by M. J. Plavec and R. K. Ulrich. Pp.598. (Reidel: 1980.) Hardback Dfl.130, $68.50; paperback Dfl.60, $31.50.

BINARY stars have always been important for astronomers because the only direct way of determining stellar masses involves the application of Kepler's laws to the observed motions of the components. In addition, a study of eclipsing binaries can give information about stellar radii and variation of brightness across a stellar disk. It has, however, become clear only fairly recently that close binary stars are involved in some of the most violent and exotic phenomena in the Universe.

These phenomena are associated with mass transfer between the components. Most stars increase in size during their life history and a member of a close binary may become so large that its surface layers are attracted to the companion star; this is called Roche lobe overflow. In addition, some stars lose mass continuously and some of this may be captured by the companion. This exchange can have important effects on the evolution of the system. If the initially more massive star becomes the less massive, an ill-matched binary can result in which it appears that the lower mass star has evolved more rapidly than the more massive, in contradiction to basic ideas about stellar evolution. The most dramatic effects occur when one of the component stars is a highly compact star, a white dwarf, a neutron star or a black hole, when a very large release of gravitational energy accompanies mass transfer. This process is important in many X-ray sources, novae and cataclysmic variables.

This book is the proceedings of a symposium held in Toronto in the summer of 1979. It contains over a hundred contributions and there are few papers which contain as many as ten pages. As a result, the non-specialist may have some difficulty in discovering what are the key questions and which are the most important papers. I am sorry that there is not a small number of lengthy reviews, since it is clear that there are some important unsolved problems which merit fuller discussion than they receive here. For example, the manner in which close binary systems are formed is not clear and there are some particularly difficult problems and sharp controversies relating to the structure of binaries that are so close that they really form a single star with two nuclei. Further, most calculations of stellar evolution involving mass transfer have assumed conservation of mass and angular momentum by the system, whereas an uncertain amount of each will escape.

Although close binaries are of wide general interest to astronomers, this book is most suitable for those who are already working in the subject.

R. J. Tayler is Professor of Astronomy at the University of Sussex, author of three elementary astronomy textbooks and Managing Editor of Monthly Notices of the Royal Astronomical Society.

Andean archaeology

John Coles

Guitarrero Cave: Early Man in the Andes. Edited by Thomas F. Lynch. Pp.328. (Academic: 1980.) $20, £13.60.

FOR some years now it has been an archaeological conceit to speak of interdisciplinary projects as the sole way of comprehending any ancient site, settlement or landscape. In few cases have these concepts been achieved, and instead the archaeological world has been presented with rather feeble attempts to coordinate and integrate the variety of natural scientific techniques that can be brought to bear upon most archaeological sites or areas. In the present book we have just such an effort, one which succeeds in part — but not wholly — in demonstrating the virtues of an interdisciplinary approach.

The cave of Guitarrero lies high (2580m) in an intermontane valley bounded by the Cordillera Blanca and C. Negra in Peru. The archaeologist will see the possibilities for an ecological research project, with dramatic seasonal variations in environmental conditions presumably being welcomed by the ancient inhabitants, and indeed this form of project is what was conceived and executed. The site was excavated, and its stratigraphical deposits dated by radiocarbon to a period which on balance seems to fall between 10,000 BP and 7,000 BP. The environmental implications and attempted reconstruction of the ancient landscapes are heavily reliant on pollen analyses, and there is relatively little on geomorphological processes which one could conceive of as affecting the landscape to a significant degree. Nonetheless, this initial section of the book sets the scene for the major thrust of the project, the study of the quite extraordinarily abundant remains of plants. This is explored through exhaustive analysis of maize and beans, with essays on the origin and early development of cultivation in highland Peru.

A third section provides a full description of the multifarious stone artifacts from the major stratigraphic phases of the cave occupation, and is somewhat tedious in approach. More interesting by virtue of rarity are the bone and wood artifacts, and remains of cordage, basketry and textiles which conditions in parts of the cave have allowed to survive. The final summary and overview of the site emphasizes this organic content which alone would have made the excavation worthwhile.

Yet, as regards disciplines, multidisciplinary it is, interdisciplinary it is not. Archaeologists may never be able to achieve the latter if they remain content with the former.

John Coles is Professor of European Prehistory at the University of Cambridge.

BOOKS RECEIVED

Mathematics

BOWEN, B.D. and WEISBERG, H.F. An Introduction to Data Analysis. Pp.213. Hbk ISBN 0-7167-1173-7; pbk ISBN 0-7167-1174-5. (W.H. Freeman: 1980.) Hbk £8.80; pbk £3.95.

ELANDT-JOHNSOM, R. and JOHNSON, N. L. Survival Models and Data Analysis. Pp.457. ISBN 0-471-03174-7. (Wiley: 1980.) £18.70.

TJUR, T. Probability Based on Radon Measures. Pp.232. ISBN 0-471-27824-6. (Wiley: 1980.) £18.50.

Physics

AGAJANIAN, A.H. (Compiled by.) Mosfet Technologies: A Comprehensive Bibliography. Pp.377. ISBN 0-306-65193-9. (Plenum: 1980.) $95.

DALTON, B.J. *et al.* (eds). Theory and Applications of Moment Methods in Many-Fermion Systems. Proceedings of the International Conference held at Iowa State University, Ames, Iowa, September 1979. Pp.511. ISBN 0-306-40463-X. (Plenum: 1980.) $59.50.

DUFFIN, W.J. Electricity and Magnetism. 3rd Edn. Pp.467. Flexi ISBN 0-07-084111-X. (McGraw-Hill: 1980.) £7.50.

KANE, J. W. and STERNHEIM, M. M. Physics. S.I. Version. Pp.664. ISBN 0-471-08036. (Wiley: 1980.) £13.80.

LANCASTER, G. DC and AC Circuits. 2nd Edn. Pp.325. Hbk ISBN 0-19-85148-X; pbk ISBN 0-19-851849-8. (Oxford University Press: 1980.) Hnk £20; pbk £9.95.

MARTON, L. and MARTON, C. (eds). Advances in Electronics and Electron Physics, Vol.54. Pp.318. ISBN 0-12-014654-1. (Academic: 1980.) $38.50.

PREPARATA, G. and AUBERT, J.-J. (eds). Probing Hadrons with Leptons. Proceedings of the 4th Seminar held in Erice, Sicily, March 1979. Pp.507. ISBN 0-306-40438-9. (Plenum: 1980.) $59.50.

RICHARD, P. (ed.) Atomic Physics Accelerators. Vol.17, Methods of Experimental Physics. Pp.641. ISBN 0-12-475959-9. (Academic: 1980.) $70.

SEPTIER, A. (ed.). Applied Charged Particle Optics. Part A. Advances in Electronics and Electron Physics Supplement 13. Pp.356. ISBN 0-12-014573-1. (Academic: 1980.) $41.

TURCHI, P. J. (ed.). Megagauss Physics and Technology. Proceedings of the 2nd International Conference held in Washington D.C., May-June 1979. Pp.683. ISBN 0-306-40461-3. (Plenum: 1980.) $69.50.

Chemistry

CRIDDLE, W. J. and ELLIS, G. P. Spectral and Chemical Characterization of Compounds. A Laboratory Handbook. 2nd Edn. Pp. 115. Hbk ISBN 0-471-27813-0; pbk ISBN 0-471-27812-2. (Wiley: 1980.) Hbk £13, $38.50; pbk £4.50.

FRANZOSINI, P. and SANESI, M. (eds). Thermodynamic and Transport Properties of Organic Salts. IUPAC Chemical Data Series, No. 28. Pp.376. ISBN 0-08-022378-8. (Pergamon: 1980.) $90.

KUWANA, T. (ed.). Physical Methods in Modern Chemical Analysis Vol.2. Pp. 411. ISBN 0-12-430802-3. (Academic: 1980.) $45.

ROSENBLATT, G. M. and WORRELL, W. L. (eds). Progress in Solid State Chemistry, Vol.12. 332. ISBN 0-08-022846-1. (Pergamon: 1980.) $34.

SAMMES, P. G. (ed.). Topics in Antibiotic Chemistry, Vol.5. Pp.312. ISBN 0-85312-191-5. (Ellis Horwood Limited: 1980.) Distributed by Wiley. £25.

SVEHLA, G. (ed.) Wilson and Wilson's Comprehensive Analytical Chemistry, Vol.10. Organic Spot Test Analysis: The History of Analytical Chemistry. Pp.282. ISBN 0-444-41859-8. (Elsevier Scientific: 1980.) $73.25, Dfl. 73.25.

VOL'PIN, M. E. (ed.). Chemistry Reviews, Vol.2 (1980). Soviet Scientific Reviews, Section B. Pp.469. ISBN 3-7186-0018-8. (Harwood: 1980.) $90.

Technology

COUNCIL ON ENVIRONMENTAL QUALITY AND THE DEPARTMENT OF STATE. BARBEY, G. O. (Study Director). The Global 2000 Report to the President of the U.S. entering the 21st Century. Vol.II: The Technical Report. Pp.766. ISBN 0-08-024618-4. (Pergamon: 1980.) £20.

DOLMAN, A. J. (ed.). Global Planning and Resource Management. Towards International Decision Making in a Divided World. Pp.262. Hbk ISBN 0-08-026309-7; pbk ISBN 0-08-026309-8. (Pergamon: 1980.) Hbk np.; pbk £3.50.

FRANCIS, W. and PETERS, M.C. Fuels and Fuel Technology. A Summarized Manual. 2nd (SI) Edn. Pp.716. Hbk ISBN 0-08-025249; flexi ISBN 0-08-025250-8. (Pergamon: 1980.) Hbk £42.50, $95; flexi £14.95, $34.50.

HASSELMAN, D.P.H. and HELLER, R.A. Thermal Stresses in Severe Environments. Proceedings of the International Conference held at Virginia Polytechnic Institute and State University, Blacksburg, March 1980. Pp.737. ISBN 0-306-40544-X. (Plenum: 1980.) $75.

HERTZBERG, R.W. and MANSON, J.A. Fatigue of Engineering Plastics. Pp.295. ISBN 0-12-343550-1. (Academic: 1980.) $35.50.

THE INSTITUTE OF CERAMICS. Health & Safety in Ceramics: A Guide for Educational Workshops & Studios. Pp.36. Flexi ISBN 0-08-026173-6. (Pergamon: 1980.) £2.

NORTHRUP, C. J. M. Jr (ed.). Scientific Basis for Nuclear Waste Management, Vol.2. Proceedings of the International Symposium, sponsored by the Materials Research Society, held in Boston, Massachusetts, November 1979. Pp.936. ISBN 0-306-40550-4. (Plenum: 1980.) $65.

WILSON, R. (ed.). Energy for the Year 2000. Pp.401. ISBN 0-306-40540-7. (Plenum: 1980.) $42.50.

ZALESKI, E. and WIENHART, H. Technology Transfer between East and West. Pp.435. Flexi ISBN 92-64-12125-0. (OECD: Château de la Muette, 2, rue André-Pascal, 75775 Cedex 16, 1980.) Np.

Earth Sciences

CARP, E. (compiled by) Directory of Wetlands of International Importance in the Western Palearctic. Pp.506. Flexi ISBN 2-88032-300-2. (UN Environment Programme and the International Union for Conservation of Nature and Natural Resources: 1980.) $27.50 + postage, from Unipub, 345 Park Avenue South, New York 10010, USA, and Bowker Publishing Company, P.O. Box 5, Epping, Essex CM16 4BU.

COPE, J.C.W. *et al.* A Correlation of Jurassic Rocks in the British Isles. Geological Society Special Report No. 14, Part 1. Pp.74. Flexi ISBN 0-632-00712-5. (Blackwell Scientific: 1980.) £8.

FROST, D.V. and SMART, J.G.O. Geology of the Country North of Derby. Memoir for 1:50,000 Geological Sheet 125. Geological Survey of Great Britain: England and Wales. Pp.200. ISBN 0-11-8841-19-X. (HMSO: 1979.) £18.

KRAFFT, D. and DRAFFT, M. Volcanoes: Earth's Awakening. Pp.160. ISBN 0-8437-3760-3, ISBN 0-8437-3761 (deluxe). (Hammond Incorporated, Maplewood, New Jersey 07040: 1980). Np.

MERRILL, R. B. (ed.). Proceedings of the Conference on The Lunar Highlands Curst, Houston, Texas, November 1979. Gosmochimica Acta, Supplement 12. Pp.505. ISBN 0-08-026304-6. (Pergamon: 1980.) $52.

RUSSELL, P. J. (ed.). Aqualine Thesaurus. Pp. 684. ISBN 0-85312-266-0. (Ellis Horwood Limited: 1980.) £45.

TURNER, F. J. Metamorphic Petrology. Mineralogical, Field, and Tectonic Aspects. 2nd Edn. Pp.524. ISBN 0-07-065501-4. (Hemisphere Publishing: 1980.) £28.50.

YEN, T. F. and WALSH, D. (eds). Energy and Resource Development of Continental Margins. Pp.238. ISBN 0-08-025127. (Pergamon: 1980.) £15.

Biological Sciences

BACHMANN, P.A. Leukaemias, Lymphomas and Papillomas: Comparative Aspects. Munich Symposia on Microbiology. Pp.273. Flexi ISBN 0-85066-213-3. (Taylor & Francis: 1980.) £9.50.

BRISTOW, A. The Sex Life of Plants: A Study of the Secrets of Reproduction. Pp.223. Pbk ISBN 0-450-04667-2. (New English Library Limited: Barnard's Inn, Holborn, London, EC1N 2JR, 1980.) £1.25.

DuPONT, H. L. and PICKERING, L. K. Infections of the Gastrointestinal Tract. Microbiology, Pathophysiology, and Clinical Features. Pp.273. ISBN 0-306-40409-5. (Plenum Medical: 1980.) $24.50.

GENEROSO, W. M., SHELBY, M. D. and de SERRES, F. J. (eds). DNA Repair and Mutagenesis in Eukaryotes. Basic Life Sciences, Vol.15. Proceedings of the Symposium sponsored by the National Institute of Environmental Health Sciences, held in Atlanta, Georgia, June 1979. Pp.458. ISBN 0-306-40552-0. (Plenum: 1980.) $49.50.

GILLOTT, C. Entomology. Pp.729. Flexi ISBN 0-306-40514-8. (Plenum: 1980.) $22.50.

HECHT, M.K., STEERE, W.C. and WALLACE, B. (eds). Evolutionary Biology, Vol.13. Pp.301. ISBN 0-306-40510-5. (Plenum: 1980.) $32.50.

KARREMAN, G. (ed.). Co-operative Phenomena in Biology. Pp.246. ISBN 0-08-023186-1. (Pergamon: 1980.) £21.25.

MAKI, A.W., DICKSON, K.L. and CAIRNS, J.C. Jr (eds). Biotransformation and the Fate of Chemicals in the Aquatic Environment. Proceedings of a Workshop held at the Univeristy of Michigan Biological Station, Pellston, Michigan, August 1979. Pp.150. ISBN 0-914826-28-X. (American Society for Microbiology: Washington DC, 1980.) $12.

MANI, M.S. and GIDDINGS, L.E. Ecology of Highlands. Monographiae Biologicae, Vol.40. Pp.236. ISBN 90-6193-093-6. (W. Junk: The Hague, 1980.) Dfl.110, $58.

RHOADS, D. C. and LUTZ, R. A. Skeletal Growth of Aquatic Organisms: Biological Records of Environmental Change. Topics in Geobiology, Vol.1. Pp.750. ISBN 0-306-40259-9. (Plenum: 1980.) $47.50.

ROODYN, D.B. Subcellular Biochemistry, Vol.7. Pp.412. ISBN 0-306-40485-0. (Plenum: 1980.) $49.50.

SANDLER, S.G., NUSBACHER, J. and SCHANFIELD, M.S. (eds). Immunobiology of the Erythrocyte. The American Red Cross 11th Annual Scientific Symposium, Washington D.C., May 1979. Pp.360. ISBN 0-8451-0043-2. (Alan R. Liss: New York, 1980.) $40, Dfl.120.

SIDDIQI, O. *et al.* (eds). Development and Neurobiology of *Drosophila*. Basic Life Sciences, Vol.16. Proceedings of the International Conference on Development and Behaviour of *Drosophila melanogaster*, held at the Tata Institute of Fundamental Research, Bombay, December 1979. Pp.496. ISBN 0-306-40559-8. (Plenum: 1980.) $49.50.

SPATZ, M. *et al.* (eds). Circulatory and Developmental Aspects of Brain Metabolism. Proceedings of the 2nd International Symposium on the Pathophysiology of Cerebral Energy Metabolism, organized and sponsored by the Serbian Academy of Sciences and Art, held in Belgrade, Yugoslavia, September 1979. Pp.446. ISBN 0-306-40542-3. (Plenum: 1980.) $49.50.

STANBURY, P. and PHIPPS, G. Australia's Animals Discovered. Pp.120. ISBN 0-08-024796-2. (Pergamon: 1980.) £23.40, £9.75.

STEVENS, D.A. (ed.). Coccidioidomycosis. A Text. Pp.279. ISBN 0-306-40410-9. (Plenum: 1980.) $28.50.

VON FABER H. and HAID, H. Endokrinologie. Biochemie und Physiologie der Hormone. Pp.174. Pbk ISBN 3-8001-2489-0. (Verlag Eugen Ulmer: Postfach 10 32 7000 Stuttgart 1, 1980.) DM 17.80.

WESTWOOD, J.C.N. The Hazard From Dangerous Exotic Diseases. Pp.223. ISBN 0-333-28360-0. (Macmillan: London. 1980.) £17.50.

WOOD, J.H. (ed.). Neurobiology of Cerebrospinal Fluid. Pp.768. ISBN 0-306-40369-2. (Plenum: 1980.) $60.50.

Applied Biological Sciences

ASSCHER, A. W. The Challenge of Urinary Tract Infections. Pp.209. ISBN 0-12-790220-1. (Academic: 1980.) £15, $34.50.

BRANDLY, C. A. and CORNELIUS, C. E. (eds). Advances in Veterinary Science and Comparative Medicine, Vol.24. Pp.325. ISBN 0-12-039224-0. (Academic: 1980.) $39.50.

BREWSTER, M. A. and NAITO, H. K. (eds). Nutritional Elements and Clinical Biochemistry. Proceedings of the 3rd Annual Meeting of the National Academy of Clinical Biochemistry, held in New Orleans, Louisiana, July 1979. Pp.463. ISBN 0-306-40569-5. (Plenum: 1980.) $45.

DAVID, G. and PRICE, W.S. (eds). Human Artificial Insemination and Semen Preservation. Proceedings of the International Symposium held in Paris, April 1979. Pp.639. ISBN 0-306-40547-4. (Plenum: 1980.) $65.

DAWKINS, M. S. Animal Suffering. The Science of Animal Welfare. Pp.150. Hbk ISBN 0-412-22580-8; pbk ISBN 0-412-22590-5. (Chapman & Hall: 1980.) Hbk £7.50; pbk £3.95.

EYSENK, H.J. The Causes and Effects of Smoking. Pp.397. ISBN 0-85117-186-9. (Maurice Temple Smith: London, 1980.) £16.50.

HARRIS, C. C., TRUMP, B. F. and STONER, G. D. (eds). Methods in Cell Biology. Prepared under the Auspices of the American Society of Cell Biology, Vol.21: Normal Human Tissue and Cell Culture, B. Endocrine, Urogenital, and Gastrointestinal Systems. Pp.512. ISBN 0-12-564140-0. (Academic: 1980.) $49.50.

HUME, D. A. and WEIDEMANN, M. J. Mitogenic Lymphocyte Transformation. A General Model for the Control of Mammalian Cell Proliferation and Differentiation. Research Monographs in Immunology, Vol. 2. Pp.251. ISBN 0-444-80218-3. (Elsevier/North-Holland Biomedical: 1980.) $83, Dfl.170.

IACOBELLI, S. *et al.* (eds). Progress in Cancer Research and Therapy, Vol.14. Hormones and Cancer. Pp.590. ISBN 0-89004-486-4. (Raven: 1980.) $49.

INGLIS, J. K. Introduction to Laboratory Animal Science and Technology. Pp.323. Hbk ISBN 0-08-023772-X; flexi ISBN 0-08-023771-1. (Pergamon: 1980.) Hbk £20; flexi £9.95.

KOLBE, H. K. (ed.) Intra-uterine Devices Abstracts: A Guide to the Literature, 1978-1979. Population Information Library, Vol.1. Pp.569. ISBN 0-306-65191-2. (IFI/Plenum: 1980.) $75.

KOLBE, H. K. (ed.) Oral Contraceptives Abstracts. A Guide to the Literature, 1977-1979. Population Information Library, Vol.2. Pp. 563. ISBN 0-306-65192-0. (IFI/Plenum: 1980.) $75.

LEVY, A. *et al.* (eds). Endogenous Peptides and Centrally Acting Drugs. Progress in Biochemical Pharmacology, Vol.16. 24th Annual OHOLO Biological Conference, Zichron Ya'acov, April 1979. Pp.160. ISBN 3-8055-0831-X. (Karger: Basel, 1980.) Sw.Fr. 82, DM 98, $49.25.

McLAREN, D.J. *Schistosoma mansoni:* The Parasite Surface in Relation to Host Immunity. Tropical Medicine Research Studies Series, 1. Pp.229. ISBN 0-471-27869-6. (Wiley: 1980.) £18.

MURPHY, E. D. and BEAMER, W. G. (eds). Biology of Ovarian Neoplasia. A Series of Workshops on the Biology of Human Cancer, Report No.11. UICC Technical Report Series, Vol.50. Pp.132. Flexi ISBN 92-9018-050-1. (International Union Against Cancer: 3 rue du Conseil-Général, -1205 Geneva, Switzerland: 1980.) SwFr. 16.

RIGTER, H. and CRABBE, J. C. (eds). Alcohol Tolerance and Dependence. Pp.455. ISBN 0-444-80212-6. (Elsevier/North-Holland Biomedical: 1980.) $95, Dfl. 195.

THOMAS, R.H. and PEREZ-MENDEZ, V. Advances in Radiation Protection and Dosimetry in Medicine. Proceedings of the Symposium held in Italy, September 1979. Pp.658. ISBN 0-306-40468-0. (Plenum: 1980.) $69.50.

WEETALL, H.H. and ROYER, G.P. (eds). Enzyme Engineering, Vol.5. Proceedings of the 5th International Conference, held in Henniker, New Hampshire, July 1979. Pp.485. ISBN 0-306-40471-0. (Plenum: 1980.) $49.50.

WINICK, M. (ed.) Nutrition and Gastroenterology. Current Concepts in Nutrition, Vol. 9. Pp.221. ISBN 0-471-08173-6. (Wiley: 1980.) £17.50.

Psychology

LIBERMAN, I.P. *et al.* Handbook of Marital Therapy: A Positive Approach to Helping Troubled Relationships. Pp.262. ISBN 0-306-40235-1. (Plenum: 1980.) $19.50.

SINZ, R. Chronopsychophysiologie. Chronobiologie und Chronomedizin. Band 217. Pp.360. (Akademie-Verlag, DDR-108 Berlin: 1980.) DM 16.

STADDON, F. E. R. (ed.). Limits to Action. The Allocation of Individual Behavior. Pp.308. ISBN 0-12-662650-2. (Academic: 1980.) £24.

Sociology

BRIERLEY, J. Children's Well-Being. Pp.171. Pbk ISBN 0-85633-218-6. (NFER Publishing Company Ltd: Darville House, 2 Oxford Road East, Windsor, Berks SL1 2DQ, 1980.) £6.75.

NEUBERGER, E. and TYSON, L. D'A. (eds). The Impact of International Economic Disturbances on the Soviet Union and Eastern Europe. Transmission and Response. Pp.493. ISBN 0-08-025102-1. (Pergamon: 1980.) £30.

RA'ANAN, U. and ROCHE, J. P. *et al.* (eds). Ethnic Resurgence in Modern Democratic States: A Multidisciplinary Approach to Human Resources and Conflict. Pp.288. ISBN 0-08-024647-8. (Pergamon: 1980.) £15.

STEWART, D. W. The Women's Movement in Community Politics in the U.S. The Role of Local Commissions on the Status of Women. Pp.146. ISBN 0-08-025971-5. (Pergamon: 1980.) £7.50.

History of Science

CHANT, C. and FAUVEL, J. (eds). Darwin to Einstein. Pp.335. Hbk ISBN 0-582-49156-8; pbk ISBN 0-582-49157-6. (Longman: 1980.) Hbk £9.95; pbk £5.65.

JACOBS, L.L. (ed.) Aspects of Vertebrate History. Essays in Honour of Edwin Harris Colbert. Pp.407. Hbk ISBN 0-89734-052-3; pbk ISBN 0-89734-053-1. (Museum of North Arizona Press: 1980.) Hbk $22.50; pbk $9.95.

Anthropology

ROBSON, J. R. K. (ed.) Food, Ecology and Culture. Readings in the Anthropology of Dietary Practices. Pp.143. ISBN 0-677-16990-9. (Gordon & Breach: 1980.) $27.50.

SCHIFFER, M. B. (ed.) Advances in Archaeological Method and Theory, Vol.3 Pp.448. ISBN 0-12-003103-5. (Academic: 1980.) $39.50.

General

APPLE, A. Megasynthesis. Pp.512. Flexi ISBN 0-9690271-0-9. (Megasynthesis: P.O. Box 343, SubS. 11, Univesity of Alberta, 1980.) $10.

AUDUBON, J.J. The Art of Audubon. The Complete Birds and Mammals. Pp.674. ISBN 0-354-04593-8. (Macdonald-Futura: 1980.) £15.95.

BERGMANN, P.G. and DE SABBATA, V. (eds). Cosmology and Gravitation: Spin, Torsion, Rotation, and Supergravity. NATO Advanced Study Institutes Series, Vol.58. Proceedings of the NATO Advanced Study Institute held at the Ettore Majorana International Centre for Scientific Culture, Erice, Italy, May 1979. Pp.510. ISBN 0-306-40478-8. (Plenum: 1980.) $65.

DADANT, C.P. First Lessons in Beekeeping. Pp.127. ISBN 0-684-16747-6. (Schribner's: 1980.) $7.95.

DOBSON, F., HASSE, L. and DAVIS, R. Air-Sea Interaction. Pp.801. ISBN 0-306-40543-1. (Plenum: 1980.) $39.50.

GRANDJEAN, E. and VIGLIANI, E. (eds). Ergonomic Aspects of Visual Display Terminals. Proceedings of the International Workshop, Milan, March 1980. Pp.300. ISBN 0-85066-211-7. (Taylor & Francis: 1980.) £18.

HULL, F. C. The Evaluation of Risk in Business Investment. Pp.177. Hbk ISBN 0-08-024075; flexi ISBN 0-08-024074-7. (Pergamon: 1980.) Hbk £12, $27; flexi £6.50, $15.

LANGEWIESCHE, M. Venedig. Geschichte und Kunst. Erelebnis einer einzigartigen Stadt. Pp.284. Flexi ISBN 3-7701-0653-9. (DuMont Buchverlag Köln: 1979.) Np.

LONG, F. A. and REPPY, F. (eds). The Genesis of New Weapons. Decision Making for Military R & D. Pp.210. ISBN 0-08-025973-1. (Pergamon: 1980.) £11.20.

MANN, W.E. and HOFFMAN, E. The Man Who Dreamed of Tomorrow. A Conceptual Biography of Wihelm Reich. Pp.295. ISBN 0-87477-143-9. (J.P. Tarcher, Inc: Los Angeles, 1980.) $12.95.

NEGANDHI, A. R. (ed.). Functioning of the Multinational Corporation: A Global Comparative Study. Pp.230. ISBN 0-08-025087-4. (Pergamon: 1980.) £9.

PRUTHI, J. S. Spices and Condiments: Chemistry, Microbiology, Technology. Advances in Food Research Supplement, 4. Pp.449. ISBN 0-12-016464-7. (Academic: 1980.) $39.50.

ROXBURGH, N. Policy Responses to Resource Depletion: The Case of Mercury. Contemporary Studies in Economic and Financial Analysis, Vol.21. Pp.219. ISBN 0-89232-093-1. (Jai Press: Greenwich, Connecticut, 1980.) $27.50.

SAUNDERS, C. T. (ed.). East and West in the Energy Squeeze. Prospects for Co-operation. East-West European Economic Interaction Workshop Papers, 5. Pp.468. ISBN 0-333-30515-9. (Macmillan: 1980.) £20.

STEVENSON, I. Cases of the Reincarnation Type, Vol.III. Twelve Cases in Lebanon and Turkey. Pp.384. ISBN 0-8139-0816-7. (The University Press of Virginia: 1980.) $25.

SURYANARAYANA, C. (ed.). Rapidly Quenched Metals: A Bibliography, 1973-1979. Pp. 278. ISBN 0-306-65194-7. (IFI/Plenum: 1980.) $75.

TURNER, J. Reckoning with the Beast. Animals, Pain, and Humanity in the Victorian Mind. Pp.190. ISBN 0-8018-2399-4. (The Johns Hopkins University Press: 1980.) $14.

CORRESPONDENCE

Continued from p. 8

most of the students and younger zoological systematists (and now even some botanists) have adopted this strategy of dealing with the empirical world in the belief that they were coming to understand something new about the hierarchical structure of nature. Of course, we should have known all along that something was amiss since most of our older colleagues from an earlier and wiser generation had told us that the spread of cladistics will have an extremely deleterious effect on Mayrian (or Simpsonian, or Darwinian) taxonomy. Well, the damage is done: their taxonomy will never be the same again — wiped out, as it were, by an epidemic of cladistics brought on by thoughtless, unhygienic and scientifically irresponsible cladists.

Some of my colleagues doubtless (in the old sense) will consider it premature for me to write such congratulatory lines about Dr Halstead on the grounds that he knows almost nothing, and appears to be unfamiliar with the literature, of cladistics. But I remind them that with Halsteadian Truth, knowledge of cladistics is irrelevant for cladistics is a theoretical matter. So I say to Halstead that there can be no serious doubt (in the new sense, for I actually seem to be getting the hang of it!) about the importance of his discovery of false doubt to systematics and evolutionary biology. Application of Halsteadian Truth suddenly reveals why all those serious doubters have questioned the usefulness of gradualism as a doctrine. This is because, poor souls, they imagined that (1) the modern doctrine of gradualism as derived from population genetics has no known empirical relationship to the hierarchy of organisms, (2) that population genetics theory was designed to rescue Darwin's theory of natural selection when it faltered because of its untestability, and (3) that the applicability of gradualism as an explanation of taxic diversity is achieved by a simple extrapolation from artificial selection experiments and the changes in gene frequencies observed in nature. I am forced also to include among the serious doubters a variety of embryologists and developmental geneticists who, before the coming of Halsteadian Truth, were thought to be rather distinguished.

The innate wisdom Halstead shows about cladistics and evolutionary theory extends to his philosophy of science and politics. On the thinnest of pretexts he tries to convince us that his certain knowledge that evolution is gradual constitutes a perfect, and necessary, refutation of Marxist doctrine. Would Halstead, I wonder, imagine that the massive extinctions that have taken place in his gradualistic world also carry sinister implications for us? The class struggle? Or did Darwin only predict the end of his own theory? Halstead, this great seeker of truth, must really be rebuked for deliberately confounding science and politics and for attributing the troublesome notions of Social Darwinism to the entirely ethical and clear-thinking scientists and educators of the British Museum.

Donn E. Rosen

Department of Ichthyology,
American Museum of Natural History,
New York

Sir — Beverly Halstead has drawn attention[1] to the ideological basis of the present exhibition policy at the Natural History Museum, and its implications. Other aspects of present policy are scarcely less disturbing. They relate to the display of the museum's collections and to the use of the building in which they are housed.

The museum's original exhibits were systematically arranged and a great number of genera and species were shown. The building in which they were housed was designed by Alfred Waterhouse (1830–1905), one of our greatest Victorian architects. It is a superb building and is listed as a Grade I Historic Building by the Department of the Environment. Waterhouse himself took great pains over the exterior and interior detailing and the whole building has an impressive unity.

The original exhibition galleries, whose plan and detailing reflected the order and hierarchy of nature, as then perceived, remained in use until the outbreak of war in 1939. The war led to closure of some galleries, and after the war nothing much happened for a long time. Museums generally were reacting against the old cases and galleries stuffed with objects, and the authorities of the British Museum (Natural History) did not seem to know what to do about the old systematic displays.

Eventually a policy seemed to emerge, exemplified by the new Fossil Mammal exhibit opened in 1972 in what was the old Fossil Mammal Gallery. In place of the dramatic long vista past mounted skeletons of fossil elephants, the continuity of the gallery was broken up and its original architectural character subdued, though some of the detail is still visible. Scientifically the new exhibit is much better than the old one, with a great deal of information on the animals and on the rocks from which they came, attractively displayed. As Halstead has pointed out[2], everyone, from a specialist to a child, can get something out of it.

So far, the museum had pursued the policy of showing a representative series from the collections which are its *raison d'etre.* In this connection it is interesting that the British Museum Act 1963 makes no reference to the public exhibition of the collections. The explanations must be that in the eighteenth century, when the original acts were framed which the 1963 Act replaces, it was self-evident that a museum existed for this purpose.

Those who hoped that Fossil Mammals heralded the modernization of other galleries were soon to be disappointed. There was an abrupt change of policy. The museum's report[3] for 1972–74 said that "Unhappily, the existing public exhibition does not match the museum's purpose . . . it has lagged far behind this century's revolution in the natural sciences. . . .it fails to give a clear picture of natural history as the study of the world in which we live with principles which lie at the groundwork of our technologies, and which are the first principles of rational living". The failure seems disputable–the original exhibits largely gave such a picture, and the rest of the sentence is gibberish. (What are the first principles of rational living?)

The report outlined a new exhibition "much larger than the old" designed to "reflect all aspects of modern biology". It gave four themes which were to be the basis for the new exhibition: Ecology; Life processes and behaviour; Evolution and diversity; and Man. The Hall of Human Biology (opened in 1977; actually a warren rather than a hall), Ecology (1978) and the new exhibits on Man and Dinosaurs criticized by Halstead, are the first fruits of the stated policy.

It is debatable to what extent the museum should set itself up to be the fountainhead of the fashionable parts of modern biological education. In the case of the new Dinosaur exhibit, the sequence of arguments presented in the exhibit is repeated almost word for word in the accompanying booklet. This exemplifies the pointlessness of the new exhibition policy, which employs a great deal of expensive equipment to put over concepts which are better expounded on paper, and which are but distantly related to the collections and function of the museum. There is no coherent relationship between the new exhibits, the collections and the building. The recent exhibits make minimal use of actual specimens (none at all in Human Biology) and are unsuited to the exhibition space provided by Waterhouse's building. Human Biology was built inside former galleries, which it conceals completely, and Ecology was set up in the former gallery of living mammals. (It has recently been moved to the Fossil Reptile Gallery, soon to be demolished, where it looks even more out of place.)

Yet for all these unpromising omens, the museum apparently does intend to show its collections. The "Evolution and diversity" exhibits will, it is stated, "contain most of the material now on show in the Museum", exemplifying 5,000 to 10,000 species. They "will continue to provide . . . the most entertaining aspect of the Museum, halls of monsters and a stuffed zoo." Now this is exactly what the Waterhouse galleries were designed to do. It seems the height of perversity not to use them for the purpose. On the contrary, it is proposed to dismantle the 1972 Fossil Mammals exhibit (an admirable exposition of evolution and diversity) and to demolish a large range of the original galleries. Surely it would be better to do the obvious thing, use the original galleries for displaying the collections, and put exhibits such as Human Biology in a communicating building outside the original structure.

It is clear that the authorities of the museum are, at best, embarassed by the splendid interiors they have inherited. They do not know what to do with them, or how to make use of them constructively. At worst they appear to regard them merely as an impediment to be concealed and, if possible, removed.

Worse than this, there appears never to have been any coherent plan for the development of the buildings and the site as a whole. Both the original building, admirably suited to its original purpose but quite inadequate for research and storage, and the site are large, but not inexhaustible. For example, the new Palaeontology building (opened 1977) is splendid and fairly well suited for its purpose of research and storage, but it has used up the last available building plot next to the main building . This piecemeal approach exactly parallels the present attitude to the use of the interiors — exhibits and other facilities being placed wherever there happens to be some space that can be occupied. The whole problem of space and buildings at the museum is in urgent need of review, preferably by an independent body since the Trustees seem to be unable to see the wood for the trees.

D.T. DONOVAN

Department of Geology,
University College London, UK

1. Halstead, L.B. *Nature* **288**, 208 (1980).
2. Halstead, L.B. *Nature* **275**, 683 (1978).
3. *Report on the British Museum (Natural History) 1972-74*, 75 (HMSO, London, 1975).

SIR—Halstead[1] has criticized the current exhibition policy of the British Museum (Natural History) on the grounds that the new dinosaur and fossil man exhibits are vehicles for the didactic presentation of a cladistic interpretation of evolution and taxonomy. He goes on to suggest that cladism and a saltationist interpretation of the history of life provide support for a Marxist view of history by progressive revolutions.

The political argument was questioned by Hughes-Games[2] and Rothman[3] and is, really, a diversion from the main issue. Marxists and creationists may find support for their ideas in cladism and punctuated equilibria, but the interpretation placed on fact or hypothesis by people with a particular axe to grind should not affect the development of primary research.

Patterson[4] presents some arguments on the validity of cladism and its relationships to evolutionary theory. This is an area that will probably continue to be discussed for many years (no doubt to the amazement and disgust of all non-participants).

However, the main issue is the exhibitions policy of the Museum. This question can be split into two: "Are the new exhibits balanced reflections of current scientific opinion?" and "Do the exhibits satisfy the public?".

The first point is answered with a resounding "no" by Halstead, and, I think, a partial "yes" by Patterson. Patterson states that "amongst scientists in the Museum there are many different viewpoints on the value and generality of cladistic methods", but this is not reflected in any exhibit yet. Is there any plan for a presentation of the views of "classical taxonomists" who regard cladism as an important analytical tool, but not as the sole aim and guiding principle of their research, or of numerical taxonomists (again with provisos as to applicability)?

The second question has not been asked yet. The new exhibition policy was started in the mid-1970s with the important aim of presenting a dynamic view of biological processes and it was stated that ". . . the scheme will be worthily carried out only if we can provide instruction and pleasure to the population as a whole . . ."[5]. Is this aim satisfied?

I will comment only on the dinosaur exhibition. This occupies the prime gallery of the Museum — the main entrance hall. Large skeletons attract the visitor and strategically placed stalls market books, postcards, badges and models. The alcoves on either side contain the teaching parts of the exhibit and show examples of contemporaries of the dinosaurs and the much-criticized presentation of cladism. However, the interested visitor will search fruitlessly for information about life-style, function and physiology of dinosaurs. My experience from presenting lectures to school children, natural history societies and adult education classes is that all these groups of laymen want to know if dinosaurs were warm-blooded, why they were so big, how they used their horns, spines and frills, how we collect them, why they died out, and so on. When I mention taxonomy there is a general glazing of eyes, fidgeting and covering of mouths. This may be an indictment of my teaching methods, but I think that it is also a reflection of what people consider interesting.

If the Public Services Department plans displays on dinosaur biology (functional anatomy, ecology, behaviour), I apologise for these remarks. If not, I don't.

MICHAEL J. BENTON

Department of Geology,
The University,
Newcastle upon Tyne, UK

1. Halstead, L.B. *Nature* **288**, 208 (1980).
2. Hughes-Games, M.J. *Nature* **288**, 430 (1980).
3. Rothman, H. *Nature* **288**, 430 (1980).
4. Patterson, C. *Nature* **288**, 430 (1980).
5. *Report of the British Museum of Natural History, 1972–74*, p.76 (Trustees of the British Museum (Natural History), 1975).

Halstead replies

SIR — In view of the seriousness of the criticisms[1] levelled against the Natural History Museum, it seems a little surprising that they failed to elicit any response from the Director or even the Head of the Public Services Department. The only letter from the museum was from a senior member of the Palaeontology Department[2], who appears to be concerned with roundly condemning the very policies which were the subject of my initial protest.

Patterson[2] insists that "cladistics is not about evolution" and that there is "no connection between cladistics and one view of the evolutionary process". The "cladistic literature" to which he refers to support these assertions comprises only papers by Platnick[3] and himself[4], and these deal with what is known as "transformed cladistics". This is most definitely not the kind of cladistics being portrayed in the public galleries of the Natural History Museum, to which Patterson appears to be resolutely opposed. The museum's Public Services Department in fact accepts the classic version of cladistics as set out originally by Hennig in *Phylogenetic Systematics*[5]. The title of his book is frequently taken as a synonym of cladistics[6], a term which refers specifically to the importance of the branching process (speciation) in phylogeny[6]. Hennig's view of the origin of new species involved the splitting of the ancestral species as a consequence of geographical isolation (allopatric speciation)[5].

The long recognized (in spite of Patterson's contrary assertion[2]) connection between cladistics and punctuated equilibria, a fundamentally "leap" or saltation view of the process of evolution, has recently been emphasized by Cracraft[7] and noted by others[8,9]. Eldridge and Gould[10] formulated the theory of "punctuated equilibria: an alternative to phyletic gradualism" by applying the concept of allopatric speciation, as used by Hennig, to the fossil record. In a subsequent review of the topic and the controversy it had engendered, Gould and Eldridge[11] contrasted the concept of gradualism "embedded in the modern history of Western cultures" with that of the "official 'state philosophy' of many socialist nations" where the "laws of change are explicitly punctuational", quoting Engels and the official Soviet Handbook of Marxism–Leninism to this effect. Their basic Marxist approach to evolution has been generally recognized as such, by *inter alia* Gray[12] and Hughes-Games[13]. The latter and Lewin[14] hold the view that the concept of punctuated equilibria is on the way to becoming the orthodoxy of the future. This being so, it is not unexpected that many of its new-found adherents may well be unaware of the philosophy behind the cause they are espousing.

The major controversy regarding the process of evolution and the fossil record concerns the notions of the origin of new species by sudden splitting (Hennig[5], Eldridge and Gould[10,11]) — the Marxist model, as against the gradualist model associated with Darwin[15], Mayr[16] and Simpson[17]. I am associated with the latter model, not from uncritically accepting the authority of Mayr and Simpson but on the basis of my own research experience.

The death of scholarship in the Natural History Museum is not in my opinion marked by the advocacy of cladistics or even Marxism in the public galleries, as Patterson[2] seems to imagine, but rather in the manner of the advocacy. The exhibits on dinosaurs and fossil man together with their accompanying books avoid any discussion of the gradual evolutionary versus the revolutionary "leap" concepts, by the simple expedient of accepting the basic assumptions of one and ignoring those of the other[18]. It is this crude partisanship which is indeed the unacceptable "strident voice of authority".

The insistence on a chosen received dogma, such as that no fossil species can be considered ancestral to any other, has led the Public Services Department into its obvious differences of opinion with the museum's Sub-Department of Anthropology over *Homo erectus*. That "there is not any serious doubt about *Homo erectus* being ancestral to *Homo sapiens*" is contrary to the rules of the cladistics as applied in the exhibit but is in fact the considered opinion of the scientific staff of the museum's own Sub-Department of Anthropology, and certainly the literature that immediately comes to hand appears to agree with this, including as it does cladistic analyses[19-24]. The most telling evidence, however, comes from within the covers of the museum's own book, *Man's Place in Evolution*[25]. I have already drawn attention[1], in this context, to the vitally important Petralona skull, the phylogenetic position of which was recently discussed in detail by Stringer, of the Natural History Museum[19] and by others[26,27]. The skull figures in the book, but curiously has been excluded from the exhibition itself. It would have served as a dramatic illustration of a transitional form exactly intermediate between *Homo erectus* and *Homo sapiens* but would at the same time have destroyed the credibility of the cladistic dogma being promoted. In the museum's public galleries, it now transpires that, unlike elsewhere in science, "ugly facts" are not to be permitted to "slay beautiful hypotheses".

The Director in his preface writes that, like the exhibit, the book "was planned with the guidance of Museum experts, particularly from the Museum's Sub-Department of Anthropology"[25]. This gives the entirely misleading impression that what is on display and published by the museum represents the considered views of the museum's own anthropologists. This is most certainly not the case and one cannot help but wonder whether the Director is actually aware of what is going on in the institution which he heads.

Finally, lest it be thought that my previous letter might be construed as "dangerous to the unfettered development of science" [28], in my estimation it is the current policies of the Public Services Department of the Natural History Museum that are already having just such an effect. I am not opposed to the public presentation of a reasoned case for cladistics, be it transformed or classical, as I have demonstrated by employing cladistic analysis in my own research[29]; nor have I any objection whatsoever to the presentation of a Marxist interpretation of the history of life, if done in the open and scholarly manner of a Stephen Jay Gould.

My objection is to a major public scientific institution, renowned internationally for its scholarship, to be seen to be abusing its authority by attempting to impose on the general public, against the scientific judgement of its own experts, controversial concepts not by argument or discussion but simply by unsubstantiated assertion.

L. B. Halstead

Departments of Geology and Zoology,
University of Reading, UK

1. Halstead, L. B. *Nature* **288**, 208 (1980).
2. Patterson, C. *Nature* **288**, 430 (1980).
3. Platnick, N. I. *Syst. Zool.* **26**, 438-442 (1977); **28**, 537-546 (1979).
4. Patterson, C. *Biologist* **27**, 234-240 (1980).
5. Hennig, W. *Phylogenetic Systematics* (Illinois University Press, Urbana, 1966, 1979).
6. Bonde, N. *Colloques int. Cent. natn. Rech. scient.* **218**, 283-324 (1975).
7. Cracraft, J. *Phylogenetic Analysis and Paleontology* (eds. Cracraft, J. & Eldridge, N.) (Columbia University Press, 1979).
8. Hull, D. L. *Syst. Zool.* **28**, 416-440 (1979); *Paleobiology* **6**, 131-136 (1980).
9. Fortey, R. *Pal. Ass. Circ.* **100**, 10-11 (1980).
10. Eldridge, N. & Gould, S. J. in *Models in Paleobiology* (ed. Schopf, T. J. M.) (Freeman, San Francisco, 1972).
11. Gould, S. J. & Eldridge, N. *Paleobiology* **3**, 225-151 (1977).
12. Gray, G. K. *New Scient.* **86**, 402 (1980).
13. Hughes-Games, M. J. *Nature* **288**, 430 (1980).
14. Lewin, R. *Science* **210** 883-887 (1980).
15. Darwin, C. *The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life* (Murray, London, 1859).
16. Mayr, E. *Systematics and the Origin of Species* (Columbia University Press, 1942); *Principles of Systematic Zoology* (McGraw-Hill, New York, 1969).
17. Simpson, G. G. *Tempo and Mode in Evolution* (Columbia University Press, 1944); *The Major Features of Evolution* (Columbia University Press, 1953); *Principles of Animal Taxonomy* (Columbia University Press, 1961).
18. Halstead, L. B. in *Marx Refuted* (eds Duncan, R. & Wilson, C.) (Pergamon, Oxford, 1981).
19. Stringer, C. *Anthropos* **7**, 81-93 (1980).
20. Bonde, N. *Major Patterns in Vertebrate Evolution* (eds Hecht, M. K., Goody, P. C. & Hecht, B. M.) (Plenum, New York, 1977).
21. Howell, F. C. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) (Harvard University Press, 1978).
22. Pilbeam, D. *The Evolution of Man* (Thames & Hudson, London, 1970).
23. Washburn, S. L. *Scient. Am.* **239**, 146-154 (1978).
24. Wood, B. *The Evolution of Early Man* (Peter Lowe, London, 1976).
25. *Man's Place in Evolution* (British Museum (Natural History) London, 1980).
26. Jelinek, F. in *Current Arguments on Early Man* (ed. Konigsson, L. -K.) (Pergamon, Oxford, 1980).
27. Wolpoff, M. H. *J. hum. Evol.* **2**, 339-358 (1980).
28. Rothman, H. *Nature* **288**, 430 (1980).
29. Halstead, L. B. in *Problems of Phylogenetic Reconstruction* (eds Joysey, K. A. & Friday, A. E.) (Academic, London, 1981).

Science and values

Sir—Between November 27 and 30 1980 I attended, as an observer, the 9th International Conference on the Unity of Sciences (ICUS) in Miami Beach. It was host to 640 participants from 85 countries. The concensus of the overwhelming majority of those who attended was that the conference, under the banner "Absolute Values and the Search for the Peace of Mankind", was a great success. Imagine, then, the consternation with which I read your editorial comment (*Nature* 27 November p.310) on ICUS "Best not to attend on Mr Moon", which launched scathing broadsides on the founder, seemingly, the aims of the conference, and, indirectly on the integrity of the participants themselves.

Science and values, naturally, reside in different worlds. One cannot understand man's spiritual quest (for liberty, truth, beauty etc.) by empirical laws nor can one restrict scientific endeavour with religious dogma. Yet who can pretend, in a world exhibiting both desperate physical need and social disintegration, that they have nothing to do with each other? This is where I consider the ICUS makes its unique contribution — as an interdisciplinary forum where scholars can freely exchange ideas developed in their own fields about the pressing needs of the world. Few would consider this an "extraneous cause". Are we to understand then, that in referring to Moon's addresses as "vacuous", and implying the term "values" to be meaningless, the editorial policy of *Nature* is to preserve the steps up the ivory tower of scientific learning unsullied by the muddy boots of morality and social responsibility?

Another point of contention is the assertion that "the participants . . . are (not) the group ideally suited for the discussion of the broad themes . . .". If eminent academics in the fields of science, philosophy and the humanities are *not*, then who is?

The conclusion of one committee chairman at the end of the Miami conference was that probably more questions had been raised than answered. I would agree with this. The aims of a body like the ICUS are not easily achieved but the challenge to the academic community to work towards the unity of the sciences can only have positive results, more than satisfying merely an intellectual appetite.

D.M. Trubshaw

London W2, UK

NSF and cryptology

Sir — In view of the extensive recent discussion of the respective roles of the National Science Foundation (NSF) and the National Security Agency (NSA) in support of cryptological research (see, for example, *Nature* 4 September, p.2) , I believe it may be useful to restate the foundation's established policy in this area.

The essential points of our policy with respect to cryptological research are these:

(1) Since mid-1977 we have routinely referred proposals with relevance to cryptology to NSA for review. We will continue to do this. The practice serves to keep NSA informed of NSF's activities in this area, and gives NSA an opportunity to make technical comments on proposals which can be useful in making funding decisions. It is not a "clearance" process; whatever comments NSA may make are advisory.

(2) NSF has long had a policy of encouraging other agencies to support basic research in areas relevant to their missions. We have specifically encouraged NSA to establish an unclassified basic research programme, and stand ready to assist that agency in this effort. We believe it is fundamentally healthy to have alternative sources of support in important areas of science, and anticipate no difficulties in maintaining close coordination between NSF and NSA.

(3) In cases in which alternative sources of support are available, we routinely encourage principal investigators to apply to such sources as well as to NSF. However, if an investigator prefers to apply only to NSF, we will consider the proposal in the usual manner, without prejudice, and reach a decision on funding using our usual criteria and peer review process.

(4) NSF does not expect that the results of the basic research which it supports will be classified, except in very rare instances. NSF does not currently have classification authority, but it has responsibility, under routine executive orders issued by both the current and previous administrations, to refer any information which it believes might require classification to the agency with appropriate subject matter interest and original classification authority. For cryptological research, that agency is NSA. The important point here is that it makes no essential difference, in terms of the likelihood of classification, whether research is supported by NSF or NSA. This policy is of long standing, and applies to all areas of research.

(5) NSF has long-established reporting requirements which allow it to meet its responsibility for prudent use of public funds. These might not be adequate in all cases where research might have special relevance to national security, and in such cases we may consider special reporting requirements. We have not done this in the past, and we may not have to do it in the future. If we did have to establish such reporting requirements, however, we would regard this not as a change in policy but simply as a change in administrative procedure necessary to apply a long-standing policy to a changed situation.

In summary, the foundation will continue to support cryptological research, will continue to coordinate such research with NSA, and will continue to encourage NSA to develop its own basic research support programme. The results of such research have not been classified in the past, and we do not expect them to be in the future, but we will ensure that our reporting requirements are adequate to allow us to meet our responsibilities with respect to possible classification. Most importantly, the foundation has a basic policy of supporting the best research it can find in all areas of science and engineering, with the fewest possible restrictions on investigators.

Donald N. Langenberg
(Acting Director)

National Science Foundation,
Washington, DC

Behind the fridge

Sir—It is highly likely that the reservoir at the back of a self-defrosting refrigerator provides an excellent breeding ground for microorganisms (*Nature* 20 November 1980, p. 208). However, your correspondent's suggestion that these organisms might cause pulmonary infections such as Legionnaires' disease neglects to consider that they would first have to be dispersed in the form of an aerosol. It is difficult to imagine how this might happen other than when the machine is subject to violent movement.

P.A. Jenkins

Public Health Laboratory Service,
Mycobacterium Reference Unit, Cardiff, UK

Old *Nature*'s almanac for 1981

In early times the beginning of each year was marked by the publication of almanacs intended to chronicle events in the succeeding months. Most of these documents were necessarily blends of fact and fiction. The following, however, is an authentic chronicle of the new year.

January: Bitter weather. President Ronald Reagan (20 January) says at inauguration that "we have nothing to hope for but hope itself", and commits the United States to "making the shuttle work within this decade". Dr William Nirenberg appointed chairman of reconstituted President's Science Advisory Committee. Yugoslavia establishes biotechnology enterprise (capital $20 million) with licence to manufacture insulin from "porcine pancreatic tissue". *Nature* publishes 25-page human mitochondrial DNA sequence. Earthquakes in Tashkent. Mercury and Mars ill-placed for observation.

February: Bitter weather (except in Southern Hemisphere). General Alexander Haig reaffirms opposition to Salt II at Senate confirmation hearings on nomination as US Secretary of State. Mass of neutrino measured at Savannah River reactor experiment as 30±35 electron volts; US Department of Energy announces "riddle of Universe all but solved" and President Reagan gives new go-ahead for Clinch River fast reactor. *Nature* publishes first instalment of DNA sequence of human chromosome 7 (F.Sanger); Unesco announces "emergency meeting" of journals editors (Addis Ababa, 1985) to consider "sequence pollution". Interim report of Swinnerton-Dyer committee on organization of University of London says that university must save money "or face prospect of radical change". Earthquakes in Hellenic arc. Venus and Mars ill-placed for observation.

March: Bitter weather; much flood. American journal *Science* begins publication of nucleotide sequence of human chromosome 7 (W. Gilbert); Senator William Proxmire announces public hearings on question "Are sequences necessary?". US General Accounting Office recommends that Shuttle Tile Repair Module be adapted as skyscraper window-cleaning machine. Neutrino mass estimated as 50±55 electron volts by observation of diffuse ultraviolet background; NASA announces "mystery of Universe now cleared up". British universities told that in future, *all* students must pay economic fees. General Alexander Haig (at continued Senate hearing) reaffirms opposition to Comprehensive Test-Ban Treaty. Earthquakes in Caribbean. Mercury, Venus and Mars ill-placed for observation.

April: Sakharov writes to announce arrival of Spring in Gor'kii. *Nature* begins publishing errata to Sanger chromosome 7 sequence. Tanzanian Secretary-General of Addis Ababa conference on sequences calls for extension of terms of reference to include measurement of neutrino mass, and charges that present practices of journal editors discriminate against developing nations. First shuttle flight crew sues NASA over early retirement terms. British Chancellor of Exchequer levies Value Added Tax on university fees, but affirms "monetarism is working". Professor Milton Friedman appointed Companion of Honour. Earthquakes in the Levant. Mercury, Venus and Mars ill-placed for observation.

May: Unseasonable weather. Soviet government accedes to Sakharov's request to be put on trial, but postpones formal charges. Dr Edward Parkes (chairman, UK University Grants Committee) says "no compelling reason" why British universities should be closed, but confesses anxiety about recurrent grant. Centenary celebrations of British Museum (Natural History) pass off without incident; cladists and opponents silenced by Royal presence. Dr F. Sanger forswears nucleotide sequencing until "somebody decides what to do with the data". President Reagan, by executive order, merges Clinch River and shuttle projects, offering management contract to Government of France. General Haig tells Senate confirmation committee that US Administration will "seek to repeal" Anti-Proliferation Act. Earthquakes in Andes. Mars ill-placed for observation.

June: Bitter weather in Antarctica; much flood elsewhere. Merrison committee on "dual-support system" for UK academic research argues that because academic research has continued even though the system has "long since" broken down, there is no special need of new mechanisms for research support. Proxmire Senate committee condemns sequencing research as "pork-barrel biochemistry" and proposes legislation denying federal funds to institutions defining sequences of "more than 52 contiguous nucleotides". General Haig confirmed as US Secretary of State after promising to press for early signature of Comprehensive Test-Ban Treaty and ratification of "upgraded Salt II". Rockwell International (main shuttle contractor) announces development of Shuttle Tile Repair Kit as personal space transport module. Earthquakes in Italy. Mercury and Mars ill-placed for observation.

July: Inclement weather. House of Lords committee on "Science and Government" urges reorganization of UK science with separation of research councils from Department of Education, appointment of minister with reponsibility for science and re-creation of Advisory Council on Scientific Policy. UK biotechnology company Celltech takes licence on Yugloslav insulin process. Committee of Vice-Chancellors and Principals announces "agreement in principle" with British Tourist Board that universities will put facilities and staffs at disposal of tourist industry during future extended long vacations in exchange for commission on tourist receipts; growing anxiety about future of Salford and Aston universities. NIH reaches compromise with Proxmire committee on safeguards procedure for regulation of sequencing work. Eclipse of Sun (31 July) visible in Middle East and central Asia. Mercury ill-placed for observation.

August: Improbable weather. NASA commissions Boeing design of reusable long-distance transport system based on 747 with "orbital operation development capability". US Department of Energy dedicates Clinch River site as "national center" for measurement of neutrino mass. Second Voyager approaches Saturn; Jet Propulsion Laboratory announces that "mankind has learned more about the Solar System in the past month than in the previous twelve". *Nature* finishes publishing errata to Sanger chromosome 7 sequence (see April). Europe on vacation. Mercury, Mars and Saturn ill-placed for observation.

September: Bitter weather. Sakharov trial postponed. British universities postpone *sine die* opening of new academic year in plan to avoid shutdown. UK government abolishes posts of chief scientist in government departments, but denies abandonment of Rothschild "customer–contractor" principle, pleading only "further steps in battle against public expenditure". Legality of NIH Advisory Committee on Nucleotide Sequences tested in US Supreme Court; Senator Proxmire agrees to modify bill for fear that "United States will lose out to foreign sequencers". Neutrino mass held "indistinguishable from zero" on basis of tritium decay measurement; Clinch River site re-dedicated as "National Nucleotide Sequence Repository". Mercury, Jupiter and Saturn ill-placed for observation.

October: Bitter weather. Interim shuttle modification programme approved by NASA; US General Accounting Office complains that the Orbitwise Re-entry Module Assembly Kit will leave no room in the cargo bay for payload, and that the decision that the shuttle vehicle itself should be disposable will have "negative economic consequences except in California". UK research councils say in annual reports that they have survived better than feared in worsening economic climate and that abandonment of grant-making is only temporary. Dr J. L. Gowans, Secretary of Medical Research Council, denies at press conference statement attributed to him that "medical research is mostly good luck" (see *Nature* 1/8 January, p.2); rather, says Gowans, "inspired guesswork". Mercury, Jupiter and Saturn ill-placed for observations.

November: Extremes of climate. Nobel prizes: Medicine, Dr César Milstein (for monoclonal antibodies); Chemistry, Dr Fred Sanger (for a method of sequencing carbohydrate macromolecules); Physics, Professor F. W. Reines (for drawing attention to the problem of the neutrino mass). NIH announce further relaxation of sequencing guidelines; joint NIH-EMBO sequence bank commissions advanced computer design from Rockwell International. No planets ill-placed for observation.

December: Weather mixed. Swinnerton–Dyer report on organization of University of London offers "stark" choice between expulsion of and management by Imperial College; university Senate resolves that changes of function unwelcome to those affected should not be effected. Vice-chancellors announce British universities will begin academic year once cold weather is ended (to avoid excessive heating bills) and "certainly no later than beginning of tourist season". Boeing claims "synchronous orbit capability" for 747; proposes development programme for returning flying machines to Earth. Proxmire committee begins investigation of "neutrino mass boondoggle". Harvard University converts P3 laboratory to squash-court. Eathquakes in Tashkent. Mercury again ill-placed for observation.

nature

15 January 1981

4/1/82.

Too soon to build laser weapons systems

The belief that science fiction must inexorably be transmuted into fact appears partly to account for the wave of interest in laser weapons which has recently been sweeping the United States. Ever since Darth Varda (the villain in *Star Wars*) showed how easily opponents could be evaporated with a flick of a ray gun, it has been a question only of time before the Pentagon and its counterparts elsewhere should be discovered to be working on weapons systems based on lasers. In reality, the Pentagon has been spending about $200 million a year for the past few years on the technology of high-energy lasers, and development may accelerate (see page 111). Some kind of laser system is likely to be one of the first military payloads to be carried aloft by the shuttle system. And there is an expectation (not yet falsified) that the new president will be more willing to put money into novel weapons systems, even harebrained devices, than his immediate predecessors. It is no wonder that laser weapons seem like weapons systems whose time has come.

A measure of scepticism is therefore in order, and has fortunately been provided in the past few weeks by a document (*High energy laser weapons — a technical assessment*) by M. Callaham and K. Tsipis of the Massachusetts Institute of Technology. The document is no more than a review of how laser weapons might be used and of the problems that will have to be solved before practical weapons systems could be put to work. Would-be Darth Vardas may regard it as a promise of goodies round the corner. More sober reading of the text suggests that even if the formidable technical problems which are all too apparent can be solved, there will persist serious physical limitations of what can be accomplished by laser weapons. In other words, there seems no immediate reason why anybody should fear that the already fragile framework of international relations is about to be transformed by yet another new technology.

The uses suggested for the non-existent laser weapons would gladden any Darth Varda's heart (if he, she or it has one). Offensive uses include the delivery of large amounts of energy on the heads of troops on battlefields and the destruction of other people's Earth satellites. But, on paper, laser weapons could also be used defensively for intercepting hostile missiles or at least aircraft. Laser systems can in principle be based on the ground or carried aloft, either in aircraft or satellites (and the United States has at present at least one Boeing 747 fitted out with equipment designed to study the propagation of laser beams in the upper atmosphere). In all cases, the objective is to exploit the coherence and potentially large energy intensity of a beam of laser radiation to damage objects or people. On paper, at least, directing a laser beam at chosen targets should be relatively easy, given that photons have no inertial mass. Perhaps fortunately, there are snags.

One is a snag well educated schoolboys will be aware of. Although it is customary and usually fair to suppose that a beam of laser radiation is a parallel pencil beam, Huygens's principle is not suspended for the benefit of military planners. The best-ordered laser beams spread out to some degree, and the effects are far from insignificant when people talk (as some do) of using lasers based in geosynchronous orbits for damaging objects at ranges of several tens of thousands of kilometres. In most of the applications dreamed up so far, damage is determined by the maximum intensity of the beam. At a range of 1,000 kilometres, the intensity of the infra-red beam from a carbon dioxide laser with an initial aperture of a metre would have declined by a factor of 400 even in transmission through a vacuum. Propagation through the atmosphere entails more rapid degradation as a result of Rayleigh scattering by molecules, by turbulence and by absorption and scattering by water droplets or particles of dust. Operations below the stratosphere will be impeded by haze and clouds (although it is entirely feasible that laser beams could be made powerful enough to burn through some meteorological obstacles). But even in a clear atmosphere, the beams of high-powered lasers are degraded for reasons rarely encountered in the laboratory. Inevitably, there is a temperature gradient and thus a gradient of refractive index across an intense laser beam and the result is, predictably, that the beam is further broadened. The chances are that a good part of the Pentagon's expenditure on military laser systems is spent on the study of this phenomenon of "blooming". The obvious remedy is to work with ever shorter pulses of radiation, but there are limits to what can be done — if the power transmitted is to remain unchanged, there is a risk that increasing field strengths will turn the atmosphere into a plasma and thus render it opaque, or generate out of the surface layers of the target a similarly protective plasma skin.

Other physical obstacles to the development of laser weapons, problems of energy supply not least among them, promise to be just as daunting. Directors of science fiction films do not fully appreciate that lasers, like other sources of energy, themselves consume power. Even supposing ideal efficiency within a laser itself, which supposes that problems of heat dissipation within the optical system are somehow solved, the quantities of energy that must be beamed about the sky in the pursuit of military targets are certain to be huge. If, for example, the objective is to deliver something like 100 joules of energy to each square centimetre of the surface of a target at a range of 1,000 kilometres, the output of a laer would have to be some 400 megajoules, allowing for the spreading of the beam from a one-metre mirror. On paper, a few kilogrammes of hydrogen fuel (and eight times as much oxygen for combustion) might be sufficient for a single attack of such a kind. Such a rate of energy consumption might be tolerable for a ground-based weapon but would be entirely out of court, at least with present technology, for a system based in some orbit around the Earth. It is therefore not surprising that the calculations that Callaham and Tsipis have carried out, admittedly back of an envelope stuff, suggest that laser weapons systems might have masses of several thousands of tons. No doubt, by ringing the changes on pulse length, mirror diameter and the like, the people at the Pentagon would be able to arrive at better compromises, but there seems no doubt that laser weapons will be huge constructions, a little like particle accelerators in the complexity of their power supply.

Such obstacles are unavoidable. Others may be more susceptible to research and development but remain, as things are, quite overwhelming. One curious aspect of the recent excitement about laser weapons is the continuing uncertainty about the best way of using a laser beam to destroy a target. The simple notion that its surface might be melted or even vapourized entails that large quantities of energy should be delivered to the target. The mechanical shock caused by a pulsed laser beam may be a more economical way of causing damage, but is less predictable. This is clearly one area in which the military have ambitions for research and development. The technology of picking out targets at which to aim is another. Laser beams themselves may have no inertia, but the mirrors which fashion them are material objects that must be pointed in a direction

0028-0836/81/020109-02$01.00

which is likely to intercept the intended target. Given the kinds of targets the planners have in mind, however, the computational and mechanical tasks that they must solve are formidable.

This is why the issue that will face the new Administration in Washington, and its counterparts elsewhere, is whether the game is worth the candle. For whatever virtues laser weapons may eventually be shown to have, there is no doubt that they will be uncommonly susceptible to relatively simple countermeasures. A simple false skin on every warhead would do the trick, while laser weapons would themselves be put off their stride by much less powerful hostile lasers. These are the reasons why even hawks in Washington are now asking that Mr Reagan should not decide to spend even larger sums of money than at present on the development of laser weapons. For the time being at least, there are better ways in which the Pentagon could spend large sums of money. The time for going all out for laser weapons is not now but later, perhaps when somebody has a clear idea for a laser weapon system that could perform a tangible if limited role in the foreseeable future. Mr Reagan's opponents on laser weapons cannot, however, have it both ways. If the technical problems are as formidable as they say, and the opportunities as negligible, there is little danger that wasting money in this way would further disturb arms control negotiations with the Soviet Union. If there were a prospect that, say, a workable anti-ballistic missile system based on a laser might be developed within a decade, it would be a different matter. On present form, that danger is remote.

What if Britain wants not to change?

The Centre for Technical Change has made a brave start on a daunting problem — the diagnosis and, with luck, the cure of Britain's chronic problem of turning research and development into prosperity (see page 114). Sir Bruce Williams, the newly appointed director of the centre, is a prize catch. His appointment should go a long way to still some of the muttering about the centre's chances of success. Even so the centre's wellwishers will keep their fingers crossed. British institutions, but especially the civil service, habitually resist gratuitous advice, which is why organizations like the Brookings Institution have not flourished in the United Kingdom. The most obdurate of the problems which the new centre will have to solve is that of making people listen to what it has to say. To do this, it will need guile but also an imaginative, even a radical, agenda. Where should it begin?

There is no shortage of conventional accounts of what is wrong with the British economy, and with the linkage between the economy and technology. For the past thirty years, public support for research and development has been generous, even profligate, but only in retrospect has it dawned on people that much of this effort has been misplaced. Defence research, especially in aerospace and electronics, has taken the largest share of public funds. Throughout this period it has been assumed, in the face of experience and against reasonable expectations, that work carried out in secret in government laboratories would somehow filter through to revivify the civil economy. In retrospect, nobody is surprised that this has not happened, that the British aircraft industry has an annual turnover smaller than that of Boeing or that the British government is having to spend public money on last-minute attempts to foster the design and manufacture of microprocessors. Yet the huge defence laboratories remain more or less intact, ageing slowly as their staffs approach retirement.

During the past thirty years, British governments have inconsistently tried to remedy this state of affairs. They have encouraged the production of graduates in science and technology (1955–65) in the belief that such people would make a market for their services. They have encouraged the creation of large technical companies (1964–70) in the vain hope that the conglomerates would seize the opportunities which technology presents. They have wasted money and, worse, people's talents on glamorous projects with no economic benefit such as the Concorde project (inherently uneconomic) and nuclear power (where even successful innovations have been pigeonholed until recently). Throughout this period, governments have sought without conspicuous success ways of encouraging technical innovation in the general run of British industry, finding only that the small and medium-sized firms best placed to profit from radical innovation are the least able to raise the necessary capital.

This depressing record shows that there is no point in hoping that research and development, or skilled manpower, can by themselves improve the exploitation of innovation by British industry. Indeed, the chances that such an old-fashioned recipe would work are less than they have ever been. Industry, especially small industry, does not lightly think of spending extra on research and development at times like these. (The exceptions are those companies, many of them in the chemical industry, which have learned in the past thirty years that research and development is not merely profitable but indispensable.) Moreover, there is little hope that the new centre will be able to produce such a more penetrating analysis than other analysts of the ills of the British economic system that the scales will fall from people's eyes, expectations of prosperity will be reduced to some amount commensurate with the Gross Domestic Product and the ills of inflation and deflation will be cured less painfully than by monetarism.

So where is the Centre for Technical Change to look for an agenda if these obvious fields of enquiry are likely to be ineffectual? Fortunately, not every hopeful door is shut. There are two particular fields in which the centre could do sterling service. First is the British education system, now most of all distinguished by its eccentric and inexplicable differences from other educational systems in other industrialized states. British first degree courses are shorter than elsewhere, but those who embark on them have to demonstrate in advance that they have already covered much of the ground. The result is that too many narrowly educated people, mock-specialists before their time, are turned out into a world in which their teachers are constantly declaiming the need for continuing education to meet changing needs. Throughout the educational system, the pursuit of academic excellence is properly regarded as a virtue. Improperly, the use of the educational system for the pursuit of something else, say wealth or even happiness, is regarded as a kind of sin. Occasionally, out of subliminal recognition that there may be an inconsistency somewhere, people like Sir Monty Finniston are persuaded to preside over inquiries into the parlous condition of this or that profession (engineering in his case). Perhaps they are lucky, since their inquiries are always so confined that governments usually decide not to act on their recommendations or (as in Sir Monty's case) to act halfheartedly. For a nation as preoccupied as the British with education, it is increasingly ridiculous that so much of what is said should make so little sense.

The second field in which the Centre for Technical Change should take an interest is closely related. British civility is, of course, renowned, and is probably also a product of the educational system. People do not openly seek to embarrass one another. The result is that casual acquaintances need not fear that they will have to disclose their salaries at dinner parties, which makes for congeniality. Unfortunately, the same principle applies in the real world. People in industrial research make their reputations by publishing in academic journals; occasionally they become (unpaid) visiting professors at a neighbouring university. Salesmen sell first to their sales managers, and afterwards to their customers. Rarely is it allowed that the pursuit of pound notes can be a seemly occupation — or that the consequences of success can ever be worthwhile. The problem for the centre, in tackling such a subject, is twofold. Is it capable of objective analysis and presentation? And would such an analysis, however rigorous, merely persuade the genteel British that things are better as they are?

0028-0836/81/020110-01$01.00

Reagan assailed by laser weapon lobbies

Hawks say go: doves argue caution

Washington

An early confrontation seems to be brewing between the Administration of incoming president Mr Ronald Reagan and members of the scientific community concerned with international arms control over plans to develop laser weapons for use in space.

Key congressional Republicans are said to have been told by members of Mr Reagan's transition team for the Defense Department that he wants to increase research and development on the use of lasers to destroy incoming ballistic missiles. One report claims that an early launch of the National Aeronautics and Space Administration's space shuttle will test an aiming device for a space-based laser, a project code-named Talon Gold.

But these rumoured plans are coming under strong attack from scientists who claim that laser weapons systems would not only be extremely costly and of dubious technical value, but that an accelerated research programme could upset the chances of negotiating an agreement over how such weapons should be controlled.

Speaking at last week's meeting of the American Association for the Advancement of Science (AAAS) in Toronto, for example, Dr Richard Garwin, chief scientist at IBM's Thomas Watson Research Center, claimed that excessive emphasis on directed-energy space weapons diverts funds and talent from more serious and more effective threats.

Mr Reagan's quoted views are compatible with a report from the science and technology subcommittee of the Senate Commerce Committee, which says that if rumours of an impending commitment by the Soviet Union to develop laser weapons are to be taken seriously, then there should be a significant increase in support for the Defense Department's research programme in the area.

The Senate committee admits that there is an "unsettling" diversity of opinion within the US scientific community about the technical and economic feasibility of high-energy lasers in a wide variety of potential applications, including defence, but it points out that the present Department of Defense budget allocations for high-energy laser research are not sufficient to achieve operational weapons as early as would otherwise seem possible. An increase in funds and commitment by the department would, it says, "significantly accelerate the achievement of directed energy weapons systems", a view apparently shared by leading members of Mr Reagan's defence transition team.

The Senate report provides some insight into the controversy about US efforts in laser weapons research that has simmered within the Pentagon for several years. The controversy surfaced in 1979 with reports in *Aviation Week* that the Soviet Union had begun to build a large high-energy laser at Sary-Shagan, a weapons testing area near the Chinese border, and that this appeared to be the first step towards a full-blown laser weapons system able to destroy both satellites and incoming missiles.

Defence officials have been divided on how to respond. Some — in particular, Dr William Perry, Under-Secretary of Defense for Research and Technology — have played down the potential threat, arguing that the gap between theory and practice was so large that the Soviet Union was unlikely to have a weapons system in operation until the 1990s, and rejecting claims that the Soviet Union had a significant strategic advantage in this area.

Others in the Pentagon have been dissatisfied with this reaction, and have been pushing for an aggressive laser weapons research programme, demanding substantial increases over the current $215 million a year being devoted to the high-energy laser research, primarily in the Air Force and by the Defense Advanced Projects Agency.

The dominant thinking within the Carter Administration has been that the main emphasis should remain on research. The report of the Senate committee — apparently now in agreement with Mr Reagan's advisers — is that research alone is not enough, and that "a balance between technology development and weapons system development" is needed.

The committee suggests, for example, that a special office be created within the Department of Defense to manage and direct the overall high-energy laser programme of the department, rather than distributing its research and development efforts across the different services. And it endorsed a bill introduced by Senator Howell Heflin of Alabama to create a National Laser Institute to coordinate all laser research and development.

Fears that such enthusiasm may be allowed to flourish relatively unbridled under the new administration are prompting the members of the arms control community to argue strongly that the national security implications of even an accelerated research programme should be carefully considered.

In a recent report from the department of physics at the Massachusetts Institute of Technology, Kosta Tsipis and Michael Callaham point out that, over and above the severe technical and economic hurdles facing the development of a full laser weapons system, even a research programme could be upsetting to strategic arms limitation talks, since it would be difficult to tell whether the research was for defensive or offensive purposes.

In his paper to AAAS, Dr Garwin argued in a similar vein that space-based lasers would be highly vulnerable to attack by anti-satellite rocket vehicles, which could be developed with a fraction of the resources required to install the space weapons.

David Dickson

European science portfolios still for grabs

Brussels

The division of portfolios among the new members of the European Commission last week has left scientific research in the dark. The next few days should give the *cabinets* of the principal commissioners, Etienne Davignon and Ivor Richard, the new British commissioner, a chance to sort out their division of responsibilities.

So far it seems that Davignon, a Belgian viscount, will be taking over responsibility for all energy and industrially related matters, while Ivor Richard will take over the rest of research, science and education. This will involve some reshuffling of the bureaucracy, but nobody is prepared to predict what effect these changes at the top will have on policy.

Davignon, now entering his second four-year stint as commissioner in charge of industrial affairs, has become one of the most powerful figures in Brussels through his handling of the steel crisis. He takes over responsibility for energy from the now departed German, Guido Brunner, and has already made it clear that he aims at a tough energy-orientated policy to revive European industry.

His chief interest in research is likely to be the development of energy-efficient technology, alternative energy sources and research aligned with industrial innovation. A notable advocate of Community programmes in telematics (the marriage of chip technology and telecommunications) and bioengineering, it is doubtful whether, with the broadening of his responsibilities, he will have much time to devote to the poorer aspects of Community research programmes.

This leaves Ivor Richard in charge of the joint research centres, education and training, which will tie in neatly with his responsibility for employment and social affairs. Another new commissioner, Karl-Heinz Narjes from the right-wing German CDU party, has been given undefined portfolios for industrial innovation, nuclear safety and environment and consumer pro-

0028-0836/81/020111-02$01.00

tection. Narjes, like Davignon a lawyer by training, is regarded as being for nuclear power and, with little in his past to suggest that he will be particularly radical in his approach to environmental issues, his coming is regarded with some apprehension.

Ivor Richard's past experience in the Labour government includes the post of parliamentary under-secretary of state for defence in 1969–70 and deputy spokesman on foreign affairs between 1972 and 1974. He was the British ambassador to the United Nations from 1974 to 1979. Previously he had been a Council of Europe delegate and, during the same period, 1965–69, had been with the West European Union.

The new commissioners as a whole are largely right-wing men with legal backgrounds. They include nobody with previous experience in scientific affairs. It may be an advantage, however, that Richard, without any other major responsibilities, will be in a position to give his full attention to promoting the interests of his directorates-general. Davignon is certainly keen to increase the Community's influence on the industrial research programmes of the member states. It will clearly be some time before the new Commission settles down and begins to exert an appreciable influence on Community policy; undoubtedly the immediate problems associated with the budget and the form of the Common Agricultural Policy will take precedence for the time being. **Jasper Becker**

Nucleotide sequences

Too many banks?

The US National Institutes of Health (NIH) and the European Molecular Biology Organization (EMBO) seem well on the way to founding a computerized bank of nucleotide sequences, but would-be competitors have not yet withdrawn from the field. The joint NIH/EMBO plan springs from meetings in Bethesda (Maryland) and Heidelberg last year, both of which acknowledge that nucleotide sequencing techniques now in use are likely to generate more information than can be assimilated informally.

For the time being, nucleotide sequences are being collected at EMBO headquarters in Heidelberg by Mr Greg Hamm. Other centres interested in the collection of nucleotide sequences include the George Washington University, Washington DC, where Dr Margaret Dayhoff plans to broaden the scope of her collection of amino acid sequences; the Los Alamos Scientific Laboratory; and Professor Richard Grantham's CNRS *Laboratoire de Biometrie* at Lyon.

The work now being carried out at Heidelberg is regarded as a necessary preparation for an eventual long-term decision by EMBO and NIH. It is by no means certain that the two institutions will continue to collaborate, although such an arrangement is seen as a potentially valuable precedent. In the meantime, the Heidelberg collection is being compiled and made available informally.

The outstanding technical problems are more concerned with molecular biology than with computer technology. Storing, retrieving and comparing nucleotide sequences stored in computer language can be tackled with existing programs, but decisions have yet to be made about the criteria by which sequences should be considered authentic, while the problems of merging incomplete sequences with each other, and of accommodating and authenticating variations, will entail something like a critical evaluation of data in molecular biology.

The NIH/EMBO plan assumes that nucleotide sequences will eventually be made available to all suitably qualified scientists, but also supposes that the effective management of a sequence repository will require the continuing presence of a small group concerned with the analysis of sequence data.

Professor Richard Grantham's plans are similar. His bank is said to contain all published sequences of more than 150 nucleotides, but tRNA and 5S rRNA sequences are not included because of their variability. At the end of September, the Lyon bank included 200 mRNA sequences together with 10 complete genome sequences and 100 untranslated sequences. A description of the bank and a list of its contents can be had free of charge from Université Lyon 1, 69622 Villeurbanne Cédex, France.

It is hoped that the future of the NIH/EMBO project will be decided early this year, after consultations with other interested parties. One of the potentially contentious issues is that of whether bodies such as EMBO and NIH should centralize an activity recognized to be essential but traditionally left to individual research workers relying on grants from similar bodies.

Polish pollution

Smelter shuts down

Poland's new Ecological Society has won a major triumph in its battle to halt pollution in Krakow — a ministerial promise that the Skawina aluminium works, one of Poland's two major producers of the metal, will be closed down entirely. The campaign against the Skawina works has been going on since the society was founded last September.

The post-war industrialization of Krakow, beginning with the building of the giant Lenin steel mills at Nowa Huta, has brought major pollution problems. By the end of the 1970s, Krakow, which lies in a hollow with little natural circulation of air, annually received an estimated 95,000 tonnes of dust and 70,000 tonnes of sulphur dioxide a year from Nowa Huta to the east, and 33,000 tonnes of dust and 33,000 tonnes of sulphur dioxide from the Skawina complex (a coal-fired power station and aluminium plant) to the south-west. Ecologists working on the former Royal Forest of Niepolomice some 20 km from Krakow had to recommend measures such as aerial spraying with lime to neutralize the acid rain.

Outside the technical journals and reports, little was published on the problem, although the local population could see for themselves the outpourings of smoke (new electrostatic precipitators were constantly being promised) and there were increasing fears of the health hazards of locally grown vegetables. The first real public hint that there was an even greater gas hazard than that of sulphur dioxide came in the autumn of 1979, with the revelation that gold art treasures in the Wawel Castle museum were becoming corroded.

A year later, however, the new Ecological Society, whose membership includes representatives of the "interested public" as well as experts, was demanding a "truthful report" on Krakow's monuments and environment including public health issues. The society's memorandum, published in the Krakow newspaper *Gazeta Poludniowa* of 20 November, demanded in particular that all processes emitting fluorine and hydrogen fluoride must be shut down until this hazard could be eliminated. In fact, both the Skawina works (the main source of fluorine pollution) and the Nowa Huta steel mill were already scheduled for modernization — at costs of 120 million zloty (£2 million) and 7,500 million zloty (£125 million) respectively. Preliminary work on these projects, it was promised, would begin during 1981.

But the debate continued and was taken up by the national television network. By the end of December, the Minister of Metallurgy, Zbigniew Szalajda, had agreed to a partial cut-back in production, and ordered 35 obsolete electrolysis tubs to be taken out of service. The Mayor of Krakow, Josef Gajewicz, increased the number to 160, cutting hydrogen fluoride emission by 60 per cent and aluminium production by almost half. On 5 January this year, the minister told the press that Mayor Gajewicz's action would produce a "very difficult situation" for manufacturers requiring aluminium. Ways would have to be found to step up production from Poland's other smelting plant at Konin. The ministry's plans for the future of the Skawina plant, he said, would be outlined shortly. Two days later, after a meeting with community representatives and journalists in Krakow, the minister announced that, "taking into account all the circumstances, the demands of the community and the economic aspects", the plant would be closed and the project "totally abandoned". **Vera Rich**

0028-0836/81/020112- .$01.00

Human experimentation

Indian rules

Lucknow

To help Indian investigators plan their clinical studies the Indian Council of Medical Research (ICMR) has formulated guidelines for conducting medical research on human subjects.

The guidelines have been prepared by the ICMR Central Ethical Committee whose chairman is a legal expert and whose members include the Drugs Controller of India. Briefly, they seek to ensure that the rights and welfare of subjects participating in a trial are protected, that the benefits accruing from the research outweigh the risks to the individual, that the "informed consent" of the subject is obtained before the start of the trial, and that the investigator has the competence and experience to conduct the trial.

ICMR has forwarded the guidelines to the Indian government and hopes that they will be followed by all institutions engaged in clinical research in the country. Institutions under the aegis of ICMR have already been instructed to follow the guidelines. It is suggested that, to implement the new code, each institution conducting clinical research should form its own committee, which would consist of five to seven experienced clinicians and non-medical persons able to give advice on legal and ethical matters. These committees would then evaluate each proposal for research on human subjects by following a nine-point code. ICMR recommends that, if the ethical committees are to be effective, they should be independent bodies.

Recognizing that it may take some time for each institution to form its own committee, ICMR has offered to provide the services of an ethical committee to evaluate projects during an interim period of one year. This new role for ICMR will enable Indian institutions which do not have ethical committees to apply for funds from international agencies such as the World Health Organization, which will not consider a proposal for research support unless approved by an ethical committee.

The ICMR guidelines stress the importance of obtaining the informed consent of each subject. Although the details for obtaining consent are left to individual committees to work out, ICMR insists that the subject or his guardian should be given as much information as possible about the drug to be studied, including its side effects, and that it should be made clear to him that he is free to refuse to take part in the trial.

Indian law demands that the permission of the Drugs Controller of India is obtained before a clinical trial is performed on a new drug. ICMR proposes to hold meetings with other organizations involved in clinical research, to formulate ethical safeguards that will prove generally acceptable. **Zaka Imam**

Vitamin deficiencies

Rickets at bay

Vitamin D-deficiency rickets appears to be on the decline among Britain's Asian community. The Department of Health and Social Security last week reported* a fall in the incidence of rickets among Asian children in the past ten years, but says that the disease remains a problem. Seeking to eradicate the disease, a department working party recommends vitamin D supplements for children at risk rather than the fortification of specific foods, milk or chapatti flour.

Once known as the "English disease", rickets was common in Victorian England with as many as a half of the children in poor areas in Leeds said to be suffering from it at the end of the nineteenth century. In the period 1926–42, the incidence was still high, with some 2–8 per cent of children in British cities with the condition. During the Second World War, however, rickets was almost eradicated by a programme (begun in 1943) in which all children received vitamin D in the form of cod liver oil supplements. The hope now is that the same approach could work among Asian children now at risk.

Rickets and its adult equivalent, osteomalacia, were first documented as a serious problem in the Asian community in Glasgow in the early 1960s, and since then has been observed in most of the larger Asian communities in the United Kingdom. Neither disease is notifiable, however, and data collection is not easy.

The recent decline of incidence is suggested by figures from the Hospital In-Patients Enquiry Programme and from a survey involving 130 general practitioners in areas with large immigrant populations. In 1977, in the Midlands and north of England, ten Asian children in 1,000 under the age of 16 had rickets, compared with only four per 1,000 in London. The risk factors are said to be lack of sunlight, strict vegetarian diet, poor rural background and a family history of rickets or osteomalacia. As both bone diseases are common in the Indian subcontinent, the report suggests that immigrants may bring it with them.

Thus, the department says, it should not be difficult to identify those at risk, which is why its report recommends vitamin D supplements backed up by health education. The decision against food fortification has been recommended by the department's Committee on Medical Aspects of Food Policy, which has studied the problem for four years, but will not satisfy all clinicians responsible for treating rickets in the Asian community.

The case for the more radical programme of fortifying specific target foods has fallen because legislation would be required before vitamin D could be added to specific foods such as milk or flour, and because the food policy committee has argued that there are sound medical reasons for not taking such a course. Thus there might then be excess vitamin consumption and, as a consequence, the risk of hypercalcaemia in children and possibly of cardiovascular complaints in adults.

In retrospect, the hypercalcaemia observed in young infants in Britain in the 1950s was almost certainly due to vitamin D overdosing. The department's report says that if milk were to be fortified with vitamin D, some children could well receive too much, especially if concerned parents also choose to provide vitamin supplements.

Technical problems complicate the fortification of chapatti flour — the target food best suited to the needs of the Asian community in Britain. Although vitamin D is stable in the flour, up to 50 per cent can be destroyed in cooking. Choosing the correct amount to add to flour is therefore difficult and the concentration of vitamin D needed to prevent rickets in young children might well prove a health risk to teenagers and adults. **Alastair Hay**

**Rickets and osteomalacia.* Report of the Working Party on fortification of food with vitamin D, Committee on Medical Aspects of Food Policy, Department of Health and Social Security. HMSO, £3.90.

Research development

Soliciting custom

The National Research Development Corporation (NRDC) is trying to improve its image. To that end it held a meeting in Manchester last week to spread its message among leading academics, industrialists and financiers in the north-west. The meeting was the second of its kind, the first having been held in Cambridge about a year ago, and it is unlikely to be the last.

NRDC, which exists to help exploit technological innovation in industry, clearly believes that not enough demands are being made on its funds and services and that more universities, and especially industries, should be coming to it with their ideas. It supports fewer projects in relation to the number of higher education and research establishments and industries in the north-west than in any other region — hence the choice of last week's meeting venue.

To judge from their criticisms, however, few of those involved in the business of technological innovation regard NRDC as a particularly helpful body. Many academics, who are obliged to offer NRDC first refusal of inventions resulting from publicly funded research, say that it is too bureaucratic, especially in securing legal protection for non-patentable inventions such as computer software. It is also criticized for not being sufficiently aggressive in marketing inventions, and for taking both too few and too many risks. Its high failure rate, as much as 80 per cent in some fields, can be taken either as a healthy

0028-0836/81/020113-02$01.00

sign of its willingness to take risks or as a reluctance to match its patent portfolio to its financial resources.

Most of the inventions communicated to NRDC which later result in commercial success come from government research establishments, universities and industry. In the year up to March 1980, 1,786 inventions were communicated to NRDC, 68 patents were assigned to it, 56 licence agreements were completed and 157 development projects were set up. For those working outside industry, the corporation offers to patent ideas, put up funds for further development to the point of commercial exploitation and then licence the inventions with industry.

NRDC's aims seem laudable enough, but academics in particular feel that the practice does not match the theory. Those who have been able to approach industry directly with their inventions often report greater success, largely because they are able to negotiate relatively informal deals which do away with the legal wrangling often associated with NRDC negotiations. Universities have also become increasingly aware of their science and technology departments as potential revenue earners.

For industry, however, NRDC's services may seem more attractive and it was at industry that last week's conference was primarily aimed. If academics criticize NRDC for not driving hard enough to exploit ideas, the NRDC may equally accuse academics of lacking motivation. Appealing to industry itself to come forward with its own inventions might be more productive. **Judy Redfearn**

British technology

New direction

The Centre for Technical Change, the technological think-tank set up in Britain last summer, has brought off a considerable coup by the appointment of Sir Bruce Williams (aged 62), Vice-Chancellor of the University of Sydney, as its first director. The appointment, announced last week, will take effect full-time when Sir Bruce bows out of his post at Sydney at the end of June, two months early.

The centre was set up with a grant of £1.5 million from the Leverhulme Trust and with the promise of further substantial funds (up to a total of £2.25 million over five years) from the Science and Social Science Research Councils. Its origins lie in the conviction of its sponsors that pure reason may help to explain the poor industrial performance of the United Kingdom in recent decades, especially in the exploitation of technical innovation.

One of the ironies of Sir Bruce Williams's appointment is that it is almost exactly 25 years since the appearance of a report on the same set of problems produced under the title *Science and Industry* by Williams and Dr (now also Sir) Charles Carter, recently retired Vice-Chancellor of the University of Lancaster and now director of research at the Policy Studies Institute, London. That document raised the question of why Britain, one of the principal spenders on research and development, should apparently derive so little benefit as a consequence.

Detailed planning of the centre's work is still incomplete. It has however been decided that the centre should be based in London, not (as some had hoped) in a university town.

It was also announced last week that the deputy director of the centre will be Dr A.J. Kennedy, research director of Delta Metal Company Limited. Dr Kennedy, who is 59, has a background in academic life (Cranfield) and in industrial research. The centre is managed by a council whose chairman is Lord (formerly Sir Michael) Swann, now Provost of Oriel College, Oxford. Its future is for the time being assured only for five years, but it is hoped that it will be able to recruit extra funds during that period and also to persuade its sponsors that it deserves a longer life.

Williams — back to a sinking ship?

Latin America

Freedom lost

Washington

As one of a set of new initiatives designed to help persecuted scientists in Latin America, officials of the American Association for the Advancement of Science (AAAS) intend to monitor loans provided by international development banks to research and training institutions in the subcontinent.

This action is a response to evidence presented to a workshop held during AAAS's annual meeting in Toronto last week that several loans may have been granted to Latin American institutions to compensate for the loss of scientists who have fled abroad.

There was particular concern over two loans — $66 million lent to Argentina in 1979 by the Inter-American Development Bank (part of the World Bank) to train scientific and technological research workers at nine regional institutes, and a $32.5 million loan last autumn to the University of Uruguay.

According to participants at the meeting, one of the purposes of the Uruguayan loan is to employ foreign faculty and consultants on a temporary basis. These workers would help to rebuild science programmes, particularly in agricultural research, that had disintegrated following the recent wave of political repression, when many university teachers had been imprisoned or had left.

The workshop recommended that scientific organizations should encourage international lending institutions to consider human rights when granting loans to educational and scientific institutions. They stressed the need to establish beyond doubt that scientific and academic freedom would be preserved.

The workshop's recommendations were made in the light of evidence that the repression of scientists in particular — and educationists in general — is becoming a chronic problem in many Latin American countries.

Several participants admitted that the situation in some of the countries had improved slightly in the past few years. There are fewer reports of repression from Brazil and Chile than in the early 1970s, and even in Argentina — Buenos Aires physicist Dr José Westerkamp said things were no worse than in the past.

But Dr Westerkamp added that there was still widespread fear that the repression might come back again. And AAAS's Committee on Scientific Freedom and Responsibility, which sponsored last week's workshop, has turned its attention to countries such as El Salvador and Guatemala, citing evidence of "generalized violence" against university and other teachers.

The evidence presented of a general decline of academic and scientific freedom in Latin American countries, including the dismissal of many scientists from teaching and research posts, was sufficient to persuade the AAAS Council, meeting last week to call for "the exploration of new initiatives to protect scientists".

One of these initiatives is likely to be the setting up of a Latin America Regional Centre for Human Rights — possibly in Venezuela — to monitor the state of scientific and academic freedom in the subcontinent, since several participants stressed the need to involve indigenous scientific societies.

Meanwhile AAAS officials said that the association hoped to invite officials from the Inter-American Development Bank to a meeting of its scientific freedom committee, and pointed to the success of international conservation groups in slowing down one loan because of its potential environmental consequences as evidence that bank officials seemed willing to discuss such issues. **David Dickson**

0028-0836/81/020114-01$01.00

CORRESPONDENCE

Explosive reaction

Sir — Lovins'[1] review "Nuclear weapons and power-reactor plutonium" lacks mention of prior rebuttals to his position[2,3]. Moreover, in following the transition of this article with a widely circulated draft (personal communication, 20 March 1978), one senses a *post hoc* justification for his testimony at the Windscale nuclear-fuel reprocessing hearings.

The conclusions drawn by Lovins are unsupported by the analysis. He implies that "power reactors are not an implausible but are rather potentially a peculiarly convenient type of large-scale Pu production reactor". If power reactors were such practical sources, nuclear-weapons states would not be operating special-purpose military plutonium production reactors — they would be buying up reactor-grade plutonium for their weapons programmes, instead of storing it in pools. Reactor-grade plutonium is a dangerous substance, able to sustain a nuclear explosion. But Lovins fails to quantify its military value.

For governments that contemplate becoming nuclear-weapons states, the accessibility of plutonium is only one factor: others are the quality of material, the availability of facilities, and the possibility of nuclear-explosive testing. There are numerous disincentives too.

For "amateurs", there are hurdles not mentioned by Lovins. In general, plutonium generated in power reactors is not suited for military applications, and unauthorized access has been severely restricted by physical safeguards. His superficial treatment does not help our common goal of minimizing the propagation of nuclear weapons.

There are some specific technical objections: His introductory paragraphs are not justified by the subsequent analysis. He applies such terms as "unsuitable", "usable", "inferior", and "less than optimal", to power-reactor Pu. He generates a list of "assumptions" that represent selected extremes, later to render them "false". Although he says "We must know exactly what 'less than optimal' [Pu for bomb-making] means", he does not define it. Lovins eventually admits that there is "somewhat greater technical difficulty in using power-reactor Pu for effective military bombs."

His "conservative assumption" that arbitrarily high even-isotope content will not affect his arguments has been challenged, although Lovins does not mention the contradiction. On that assumption, Lovins proposes that "Pu discharged from power reactors will. . . have major military potential". His Table 2 does not reach high concentrations of the even-plutonium isotopes, thereby preventing the reader from making independent judgements on the efficacy of isotopic denaturing at higher even-isotope concentrations. Independent estimates of dramatic reductions in explosive yield that accompany isotopic dilution[3,4] are not included in Table 2.

He categorizes his choice of parameters as "realistic", thereby dismissing the deleterious effects of other forms of denaturing. Lovins claims that my "analysis of weapons physics seems deeply flawed". That is the entire support he gives to the charge. Yet, his own "weapons physics" is sparse, lacking any equations, and sustained mainly with citations from sources which I consider discreditable.

There are many misunderstandings, oversimplifications and omissions in Lovins' treatment of near-critical fissile masses: the difference between reactivity and excess reactivity; the role of synergistic effects; and his inability to distinguish between laboratory and field developments. I have identified and studied eight phenomena that can adversely affect the yield of an explosive configuration: critical mass dilution, metallurgical phase change, radial compression, subcritical multiplication, predetonation, generation time, reactivity limit and surface leakage. Lovins ignores most of them.

The interpretation that "the least favourable moment cannot get any less favourable" for an explosive yield is contradicted by Poisson statistics — a predetonation contour that Lovins notes but fails to apply. His repeated conception that the "worst case, minimum, 'fizzle' yield is still. . . 'militarily useful' ", is inconsistent with probability theory.

Examination of the citations taken by Lovins from Wohlstetter and Gilinsky indicates the origin of his circular arguments regarding plutonium denaturing. Of the limited official information generally available, most current reviews depend heavily on R.W. Selden's unpublished conclusions, which are self-contradictory.

It is incorrect to state "*categorically* that bombs from reactor-grade or deliberately 'denatured' Pu are less effective, less powerful or less reliable than those made from weapons-grade Pu". On the other hand, using a term from Lovins, it is "disingenuous" to deny that those attributes may be applied to the average properties of denatured plutonium. I estimate that plutonium can be isotopically denatured as effectively as ^{235}U or ^{233}U at a given fissile fraction. All reactor fuels and reactor types pose comparable levels of difficulty regarding weapons application, given corresponding safeguards and fissile fractions. Weaknesses of one or the other should be identified and remedied. However, the relative reactor usability and supply of isotopically denatured plutonium is another matter.

The point that the "marginal time and money required to use civil reactors for military production are orders of magnitude less than those needed for dedicated military facilities," has an element of truth mixed in with carefree terminology. One must differentiate between base power reactors and small research reactors. Research reactors do indeed provide a hedge, a "civilian cover" for some weapons-grade plutonium production; however, power reactors are self-protected and inaccessible under an appropriate safeguards regime.

Institutional and technological safeguards can minimize diversion and proliferation. As our respective nations expand their arsenals of destructive weapons, such misplaced concern only adds to irresolution. One can transcend some of the dispute by noting that international safeguards and a ban on all testing of nuclear weapons are the most prominent and effective steps that should be taken.

A. De Volpi

Argonne National Laboratory,
Argonne, Illinois

1. Lovins, A.B. *Nature* **283**, 817-823 (1980); erratum *Nature* **284**, 190 (1980).
2. *Power Engng*, p.32 (November 1977); *Bull. atom. Scient.* **34**, 62; 59 (1978); **35**, 50; 57 (1979); *Phys. Today* p.84, (December 1978); p.83 (July 1979).
3. De Volpi, A. *Proliferation, Plutonium, and Policy: Institutional and Technological Impediments to Nuclear Weapons Propagation* (Pergamon, Oxford, 1979).
4. Sahin, S. *Nature* **287**, 578 (1980).

Role for ESF?

Sir — One aspect of its work that the European Science Foundation (ESF) could substantially expand (*Nature* 20 November 1980, pages 202, 204) is the sponsoring of scientific conferences and schools, especially in interdisciplinary areas.

The association of NATO Advanced Study Institutes and NATO Summer Schools with the NATO military pact is a matter of continuing shame to the academic community. Scientific organizations should be pressing the national governments to agree to the ESF taking over NATO's role in conference sponsorship and funding in Europe. The US side could presumably be accommodated via bilateral NSF–ESF arrangements.

Let's end the shameful military link with scholarship forthwith!

M.K. Wallis

University College,
Cardiff, UK

Theories of cancer

Sir — In reviewing my book *The Politics of Cancer*[1], Peto[2] charged that one table (Table 6.4) selectively and deliberately omitted data that would otherwise have questioned the conclusion that low dietary doses of saccharin (0.01 per cent) are carcinogenic in rats. This statement is incorrectly based on comparison of Table 6.4 with data in an appendix to the report of the Office of Technology Assessment[3]. However, the caption of Table 6.4 clearly states that it is based on another table, prepared by Melvin Reuber for use in the congressional testimony on saccharin by the Health Research Group on 21 March 1977[4]. Peto persisted in publishing this serious allegation in spite of two warnings.

Peto also appears unfamiliar with the content of the Office of Technology Assessment's report, which he cites as the basis for his charges on saccharin. This report explicitly discusses Reuber's low-dose data and his conclusion that "the increased incidence of lymphosarcoma of the thorax in rats at the 0.01 per cent . . . are highly significant in the saccharin study"[5]. Reuber also emphasizes that the carcinogenic effects of saccharin in various studies were not always dose-related.

Contrary to Peto's impressions, inversions in dose–response data are not uncommon in both experimental and epidemiological carcinogenicity studies. Reasons for such inversions include competing risks, heterogeneity in tested populations, and statistical fluctuation, particularly when dose–response curves are shallow. Peto also fails to recognize that saccharin has produced tumours in experimental animals at low as well as at relatively high doses.

Contrary to Peto's impression that "it is not surprising that so many chemicals (such as saccharin) at such (high) doses can cause cancer in animals", there is an overwhelming consensus in the qualified, independent scientific community that high-dose testing does not produce false positive results, and that this is necessary to reduce the insensitivity of carcinogenicity tests, reflecting the small number of animals tested compared with large human populations at presumptive risk[6]. It is also well recognized that carcinogenicity testing in excess of maximally tolerated doses (MTD) can produce false negatives due to competing toxicity.

Peto's charge of selective omission of the saccharin data is not his only misrepresentation. After his circulation of drafts of his review to various US experts in the autumn of 1979, Schneiderman, then Associate Director for Science Policy of the National Cancer Institute (NCI) and co-author of the government report on the importance of occupational carcinogens[7], explained in a letter to Peto that breast cancer correlates as well with Gross National Product as fat, that occupation had been ignored in studies exclusively associating lung cancer with smoking, that there have been major recent increases in lung cancer among non-smokers and that there have been "big and frightening" recent increases in cancer incidence which cannot be accounted for by smoking or other lifestyle factors.

Similarly, Upton, then Director of NCI and a co-author of the report[7], challenged the erroneous charge that a "4–5" multiplication factor was used to inflate the 1978 government estimates of cancer anticipated from occupational exposures. ". . . This simply is not true . . ."

Repeatedly, Peto dismisses as personal views statements in *The Politics of Cancer* to which he takes exception, rather than recognizing that they are based on fully referenced primary sources and without attempting to challenge these sources directly. For instance, he disparages the conclusion "that the cure rates for major cancers have not been improving much over recent years" without noting that this reflects cited NCI data. Peto refers to the "claims that cancer costs the US economy over $25 billion per year" without recognizing that such figures are derived from NCI sources. Similarly, Peto criticizes as "misleading or unbalanced" references to the term "medical–industrial complex" without attribution to its source, referenced and identified in the text (p.77), as the caption of an editorial in *The Lancet*[8].

Peto charges that the book denigrates the value of short-term carcinogenicity tests, which he asserts is a "most inexplicable error of scientific judgement". While problems of such tests, particularly the limited associations between carcinogenicity and Ames test data for compounds from a wide range of structural classes[9,10] are recognized, the book concludes (p.68) that there is a "range of useful applications" for these tests, particularly when incorporated into battery protocols.

Peto accepts that industries "delay or obstruct any hygienic measures which will cost money, . . . (and) that the scientific literature is not immune from distortion by financial interests". Peto apparently also accepts the wide range of case studies in *The Politics of Cancer*, which document a common pattern of constraints, including manipulation, distortion and destruction, in health, safety and economic data generated or interpreted by industry and its consultants. Yet he seems willing to accept studies sponsored or endorsed by industry as authority for the nearly exclusive lifestyle theory of cancer causation. Moreover, he is opposed to further regulatory controls on grounds of costs, professed "inactive conservatism", and because "for most toxic chemicals, we now have both qualitative and quantitative uncertainty about the health benefits of restriction". This seems a questionable basis for prudent public health policy.

SAMUEL S. EPSTEIN

School of Public Health,
University of Illinois at the
Medical Center, Chicago, Illinois

1. Epstein, S.S *The Politics of Cancer* (Sierra Club Books, San Francisco 1978, revised and expanded in Anchor/Doubleday, New York, 1979, to which quotations refer).
2. Peto, R. *Nature* **284**, 297-300 (1980).
3. *Cancer Testing Technology and Saccharin* U.S. Congress, Washington, D.C., June 7, 1977.
4. Wolfe, S.M. Johnson, A. *Public Citizens Heelth Research Group, Testimony Before Subcommittee on Health, House Commerce Committee Hearings on Saccharin*, March 21, 1977.
5. Reuber, M.D. *Envir. Hlth Perspect.* **25**, 173 (1978).
6. Report to the Surgeon General, USPHS, *Evaluation of Environmental Carcinogens* (*Ad Hoc* Committee on the Evaluation of Low Levels of Environmental Chemical Carcinogens, April 22, 1970).
7. Bridbord, K. *et al. Estimates of the Fraction of Cancer in the United States Related to Occupational Factors* (NCI, National Institute of Environmental Health Sciences, and National Institute for Occupational Safety and Health, 15 September, 1978).
8. *Lancet* **ii**, 1380 (1973).
9. Rinkus, S.J. & Legator, M.S. *Cancer Res.* **39**, 3289 (1979).
10. *International Program for the Evaluation of Short Term Tests for Carcinogenicity* National Toxicology Program, Dept of Health and Human Services, April, 1980.

The article "Fallacies of lifestyle cancer theories" by S.S. Epstein and J.B. Swartz begins on page 127 of this issue.

Spanish science

SIR — The publication (*Nature* 23 October 1980, p.674) of the *Manifesto* on the status of science in Spain issued by leading academics and scientists of universities and research institutes is a major event. I would like to make a few comments from the viewpoint of a scientist who has been working closely for over ten years with colleagues at Spanish universities.

This is a time of great expectations both nationally and internationally for Spanish science and engineering. Democratic changes and heightened national expectations have opened the door for a major strengthening of scientific, engineering and technological work. The largely dormant past in universities and the undeveloped governmental infrastructure for science are giving way to a healthy stirring throughout the whole system. A recent visit to Madrid confirms a vigorous self-examination of the structure of research and teaching with significant organizational changes proposed.

There is, of course, no unique set of circumstances that will guarantee motivation of faculty and students, but there generally appear to be three essential elements. Scientists must be assured of government respect for scientific work without political interference; there must be reasonable financial support and job opportunities for researchers and graduate students; and there must be adequate and respected links between government, universities and public and private research institutions. Only the first element is now firmly in place. Difficulty in providing the second element is not unique to Spain, but occurs in many industrial countries of moderate size where the number of research positions is limited. At the moment, for this reason, some of the best Spanish science and engineering graduates see their future abroad.

Academics might themselves develop research institutes within the universities directed to national needs such as high technology, engineering services, mitigation of earthquake and other natural hazards, and agricultural improvements. Although such efforts are exhausting and time-consuming, entrepreneurship to draw in industry, provincial governments and private funds is an essential part of scientific vigour. The *Manifesto* appears to address largely the central government. No doubt, realistically at this stage, the major funding for research must come from Madrid, but, in the long run, a diversity of support would appear to be the most fruitful.

The third element should be easier to achieve since it does not in itself entail much expenditure. The need to examine critically scientific goals and infrastructure related to national needs has in other countries led to the establishment of something like a Royal Commission with independent powers to probe, report and recommend. Strengthening and linking government institutes undertaking research, similar to the highly effective scientific and industrial research organizations of Australia, New Zealand and elsewhere might be considered. There appears to be a need to integrate government scientific institutes more closely with universities, given the new social climate in Spain in which protection from debilitating political agitation is not required.

From my own experience, it is crucial for scientists and engineers in Spain to organize themselves systematically so that their professional societies and academies are representative and respected. The work of such associations can clearly demonstrate to the body politic (of all parties) that among the many competing interests in a modern state, science should have a high priority on both economic and cultural grounds. Rather than an emphasis on "rights", continual demonstration is needed of the way that scientific enquiry really maintains itself in a free society. Again, from my own experience, it is important to cultivate those members of the *Cortes* who are scientists, engineers, technologists and economists. If they are to accept the responsibility of speaking for the Spanish scientific community they must have the necessary contacts and realistic proposals that can be argued effectively.

Finally, as a sympathetic outsider, I am struck by the role that Spanish science and technology might play in the developing world. Much of Central and South America lies open, one would suppose, to the establishment of close ties with Spanish scientists and institutions. The value of common language and other heritages cannot be overemphasized. Although some cultural and scientific ties have been forged, there are exciting possibilities for an expansion of teaching, scientific exchanges, bilateral scientific and technological agreements.

Through such channels, the idealism and abilities of many Spanish science graduates could find a natural place that serves not only national self-interest but also the global responsibility of all scientists. Might we not look forward to a vigorous Spanish science that belongs not only to Europe but also to the Americas? Part of the answer lies in the force of the *Manifesto* itself.

BRUCE A. BOLT

University of California,
Berkeley, California

NEWS AND VIEWS

Strife over visual cortical function

from D.M. MacKay

IT is just over 20 years since Hubel and Wiesel[1] reported that cells in the striate cortex of cat show a preference for edges or bars of luminance-contrast at specific orientations. It seemed attractive to think of such cells as edge- or line-detectors, performing the first steps towards pattern recognition by analysing the retinal image into an array of oriented line-elements. When it was discovered that the cells also respond preferentially to gratings of specific spatial frequency-bands, however, a rival theory was advanced that they were 'really' frequency-filters analysing the image into its spatial Fourier components[2].

Complementarity of spatial and frequency analysis

But is this a genuine either-or? A simple cell can be thought of as a device for sampling the spatial distribution of luminance on the retina. As such it is subject to some limitations as first elucidated by Gabor[3] in the context of communication engineering. He showed that if we want to represent a communication signal by a succession of elementary signals (samples) in the time domain, we have to accept a compromise between the 'spread' of each sample in time and in frequency. The imprecisions of our representation in frequency (Δf) and time (Δt) are related by an 'uncertainty principle': $\Delta f . \Delta t \geq \frac{1}{2}$(ref.4). For the product $\Delta f . \Delta t$ to be a minimum, the elementary samples should not have sharp cut-offs in time or frequency. The best compromise is a sine- or cosine-wave whose amplitude is modulated by a gaussian function. It was later[5] argued for analogous reasons that functions of space (such as optical images) would be most economically analysed in terms of elementary spatial functions with gaussian profiles. Four examples of such functions, adapted from Gabor[3], are shown in Fig.1a, b, with their spatial-frequency spectra in Fig.1c. This means that any linear system with a spatial response profile of the kind illustrated in Fig.1a, b is bound to show a preference for a restricted band of (spatial) frequencies. The narrower the spatial profile (Δx), the poorer will be the frequency selectivity.

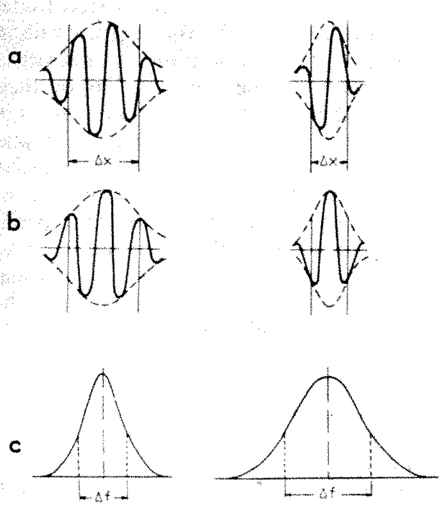


Fig.1 Analogues in the spatial domain of Gabor's sampling functions of (a) 'sine' and (b) 'cosine' type, with (c) their corresponding spatial-frequency spectra (adapted from ref.3). Note the reciprocal relation between the 'spreads' in space (Δx) and in spatial frequency (Δf).

As it happens, the receptive field profiles of simple cells (Fig.2) are remarkably similar to the right-hand examples in Fig.1 a, b. Some are 'sine' in type (Fig.2a); others 'cosine' (Fig.2b). To the extent that they respond linearly, then, simple cells cannot help showing some selectivity also in the spatial-frequency domain[6]. In communication-engineering terms, their spatial 'spread' (Δx) is of the order of X (Fig.2). Their spatial-frequency spectrum is also near-gaussian, with a 'spread' of $\Delta f(x)$ given by the 'uncertainty relation': $\Delta x . \Delta f(x) \geq \frac{1}{2}$. Taking Δx as approximately the wavelength (λ) of the sinusoidal grating that would best fit the spatial profile, this would predict a spatial-frequency bandwidth (Δf) of the order of $1/2\lambda$ — roughly one octave. As this is the order of magnitude of the bandwidths reported[7,8] for simple cells, the receptive fields of these cells seem to represent a near-optimal solution to the problem of sampling optical images in *both* the frequency and the spatial domains. Their relatively wide bandwidth (low 'Q') suggests that at this stage the inevitable information-theoretic compromise is weighted in favour of spatial resolution; but to argue whether they are 'really' concerned with one or the other would clearly be inept. The two aspects are logically complementary.

'Crucial' experiments?

In view of all this, it is surprising to find experiments that demonstrate the predictive power of Fourier analysis of stimuli still being presented as if they help decide between 'edge-detector' and 'frequency-filter' models of visual information-processing. To take a recent example, DeValois *et al.*[9] attempted "to discriminate between two models of the response characteristics of simple and complex cells of the striate cortex" by comparing responses to gratings and chessboard patterns. A chessboard has its main spatial-frequency components oriented at 45° to the edges of its squares. If the cells of striate cortex were detectors of oriented bars or edges, it was argued, they should respond best when the edges of the chessboard squares lay at their preferred orientation. If they were spatial-frequency filters, they should prefer the edges to lie at 45° to it. In the event, DeValois *et al.* found the latter result, and claimed thereby to

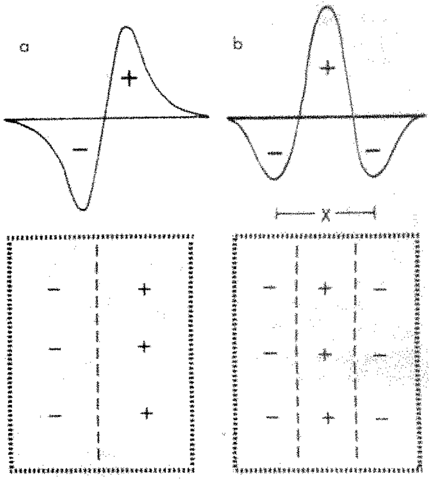


Fig.2 Typical receptive field profiles (above) and idealized field maps (below) for simple cells with (a) one 'on' (+) and one 'off' (−) zone; (b) a central 'on' (+) and two 'off' (−) zones. The ordinate denotes the contributions to the excitability (or the firing rate) of the cell made by an optimally oriented light line, as a function of distance across the field. (For cells which are silent when unstimulated, such curves can be obtained by maintaining a steady background level of activity with an auxiliary stimulus.)

0028-0836/81/020117-02$01.00

have confirmed the 'frequency-filter' hypothesis.

It is easy, however, to see how such results could be predicted without recourse to Fourier analysis. Considering first a single line of black squares (small compared with the separation of 'on' and 'off' zones) with a common diagonal (Fig.3a), everyone, including 'bar-detector' theorists, would expect the best response when the common diagonal lay parallel to the on–off boundary. Only if the squares were so large that one edge spanned the receptive field (Fig.3b) would the cell respond best when that edge was parallel to the on–off boundary. If we assume that light/dark gradients summate algebraically (even if non-linearly) over the receptive field, partial cancellation could be expected to occur in the case of Fig.3c, so that the response would be doubly position-dependent. Again, with a row of smaller squares as in Fig.3d, cancellation would be predicted[10]. With two diagonal lines of small squares (Fig.3e), Fig.2b would predict a maximal response for a critical separation (X) roughly equal to the spacing between 'off' zones — still with the diagonals parallel to the 'length' axis. With bigger squares (Fig.3f), the stimulus becomes a diagonal strip of chessboard pattern. The addition of further strips will introduce subsidiary maxima into the polar diagram at orientations giving a sufficient difference between the net excitation in the 'on' and 'off' zones. Thus none of the results presented as crucial by DeValois *et al.* requires Fourier theory for its explanation.

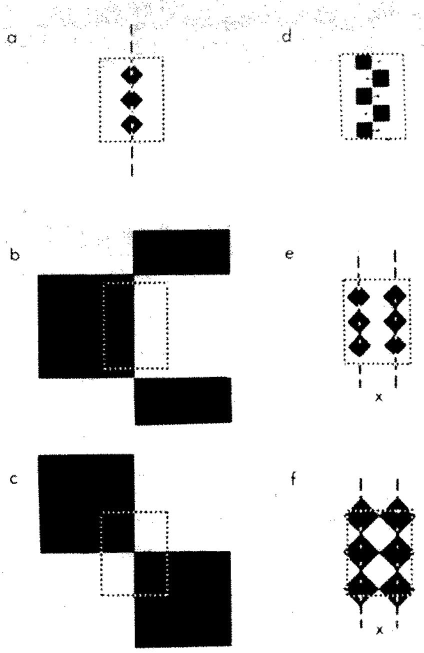


Fig.3 (a) For a line of squares small compared with the receptive field, the preferred orientation should lie along the common diagonal.
(b) For larger squares, the preferred orientation will be parallel to the edge; but
(c) If the common corner of two squares falls within the receptive field, some cancellation can be expected.
(d) A row of small black and white squares will offer oppositely oriented contributions which should cancel out.
(e) Two optimally oriented lines of small squares should be spaced by a distance X (Fig.2) to match the receptive field profile.
(f) A chessboard can be regarded as a special case of Fig.3e.

The direct calculation of net excitation for various orientations and spacings, however, is laborious and inelegant. Fortunately, a mathematical technique is available to simplify and systematize such calculations. Its name — Fourier analysis. The fact that Fourier analysis predicts the polar diagrams speaks neither for nor against the theory that the *cortex* operates as a Fourier analyser[12]. It shows only that the cell's response is (approximately) linear.

Empirical issues?

Is this then entirely a bogus issue, or are there genuine questions to be settled by experiment? Three empirical points may be noted:

Firstly, although all such cells are selective both for location and orientation and for spatial frequency, it is their relative *precision* in each domain that shows to what extent they could be useful (or useless) as filters. The frequency-selectivity of most simple cells is by engineering standards abysmal.

Secondly, the predictive success of Fourier analysis of stimuli does rule out some kinds of *non-linear* models of simple-cell function. For example, the evidence of DeValois *et al.*[9] is clearly against any idea of 'edge-detectors' as cells so non-linear as to be insensitive to the polarity of luminance edges. (Talk of oriented edges as 'trigger features' for cortical cells may have contributed to this misconception.)

Thirdly, whether the firing of any single cell *signifies* to the CNS the presence of a specific feature is an empirical question of cortical information-engineering design on which data are still scanty but thought-provoking. Hammond and MacKay[13], for example, have found that the response of even simple cells to bars can be modulated by motion of a textured background, and have shown that most striate complex cells have different polar diagrams for motion of texture and of bars[14]. Again, recent work with multielectrode recording by Gerstein and Michalski[15] suggests that more information may be represented by the pattern of temporal relationships between the firings of neighbouring units than by the firing patterns of any unit in isolation.

On both experimental and theoretical[16] grounds, then, it seems timely to reconsider the possible roles of striate cortical cells in visual pattern recognition. As individual descriptors of the stimuli falling on their fields, they are imprecise or ambiguous in all the domains of space, frequency and velocity. As members of local cooperative assemblies they could collectively offer more exact descriptions. A third possibility, however, is that their job is not to provide a symbolic *description* of the visual field at all. The task of symbolizing the perceived world could well be a more central process[17], which requires from the sensory system not pictures or descriptions but an array of selective clues to help it 'home in' on the appropriate symbolic representation. If so, we must be prepared to look for quite different kinds of link between striate cortical activity and pattern recognition.

D.M. MacKay is Head of the Department of Communication and Neuroscience, University of Keele, UK.

---

1. Hubel, D.H. & Wiesel, T.N. *J. Physiol., Lond.* **148**, 574 (1959); **160**, 106 (1962).
2. Campbell, F.W.& Robson, J.G. *J. Physiol., Lond.* **197**, 551 (1968); Blakemore, C. & Campbell, F.W. *J. Physiol., Lond.* **203**, 237 (1969).
3. Gabor, D. *J. Inst. elect. Engrs* **93, III**, 429 (1946); **94, III**, 369 (1947).
4. As a familiar example, an organ pipe has a poorly defined pitch when sounded for a time comparable with its period of vibration.
5. MacKay, D.M. *Phil. Mag.* **41**, 289 (1950).
6. This has frequently been pointed out, for example by Kulikowski, J.J. & King-Smith, E. *Vis. Res.* **13**, 1455 (1973), Appendix I.
7. King-Smith, E. & Kulikowski, J.J. *J. Physiol., Lond.* **247**, 237 (1975).
8. Andrews, B.W. & Pollen, D.A. *J. Physiol., Lond.* **287**, 163 (1979).
9. DeValois, K.K., DeValois, R.L. & Yund, E.W. *J. Physiol., Lond.* **291**, 483 (1979).
10. Experiments in collaboration with P. Hammond have shown some degree of algebraic summation of gradients for simple cells, and also for those complex cells that respond poorly to random texture. The assumption does not however hold for another class of complex cells (ref.11) which respond strongly to motion of textured fields that do not excite simple cells. Such complex cells may even respond better to a chessboard with edges *parallel* to their preferred axis than to one at 45°. Thus generalizations about the response characteristics of 'striate complex cells' cannot validly be made without specifying the subclass in question.
11. Hammond, P. & MacKay, D.M. *Expl Brain Res.* **30**, 275 (1977).
12. For analogous arguments relative to the use of chessboards as visual stimuli in psychophysics see F. Dow Smith, *Science* **198**, 207 (1977); C.W. Tyler, *Science* **198**, 208 (1977).
13. Hammond, P. & MacKay, D.M. *J. Physiol., Lond.* **284**, 117P (1978).
14. Groos, G.A., Hammond, P. & MacKay, D.M. *J. Physiol., Lond.* **260**, 47 (1976).
15. Gerstein, G.L. & Michalski, A. Paper presented at symposium on Neural Communication and Control, Debrecen, 21-24 July 1980 (in the press).
16. See also Marr, D. & Hildreth, E. *Proc. R. Soc. Lond.* B **207**, 187 (1980).
17. MacKay, D.M. in *Models for the Perception of Speech and Visual Form* (ed. Wathen-Dunn, W.) 25-43 (M.I.T. Press, Boston, 1967); *Cerebral Correlates of Conscious Experience* (eds Buser, P.A. & Rougeul-Buser, A.) 335-245 (Elsevier, 1978).

0028-0836/81/020118-01$01.00

Yeast genetics and molecular biology

from Terrance G. Cooper

LOUVAIN-LA-NEUVE was the site of the Tenth International Conference on Yeast Genetics and Molecular Biology. Twenty-four plenary presentations, eighteen evening workshops and over 200 posters allowed an exhaustive exchange of information about *Saccharomyces cerevisiae* and could lead one to conclude confidently that *S. cerevisiae* is now a favourite model for studying the structure and operation of eukaryotic cells.

RNA splicing was a major topic at the meeting and provided a focal point for studies of tRNA gene expression. The splicing system responsible for processing primary transcripts from eight or nine tRNA genes appears distinct from that processing mRNA. Mutants unable to remove intron sequences from *SUP4* tRNA primary transcripts had mutations in all portions of a recombinational fine structure map of the *SUP4* locus. This led to the conclusion that the entire three-dimensional structure of the intron-containing tRNA is probably important for recognition by the splicing enzyme system; this may act to provide a final opportunity for 'proof-reading' before activation of the newly synthesized tRNA.

Production of the var1 polypeptide (a component of the small ribosomal subunit from mitochondria) may involve yet another example of RNA–RNA splicing. A genetic element has been identified which in some way specifies the size of a transcript, but is not contiguous with the DNA coding for it. The var1 polypeptide is found in polymorphic forms ranging from 40,000 to 44,000 in molecular weight. A *var1* determinant, shown to be responsible for the polymorphism, is situated between the *ery* and *oli1* loci on the mitochondrial genome. However, complementation studies demonstrated that the *var1* determinant acts in *trans*. DNA sequencing studies revealed that the determinant does not possess apparent coding capacity (it is AT rich with no open reading frames) and is far too small to code for the var1 transcript. Normal amounts of the var1 transcript are found in petites containing only repeated copies of the *var1* determinants (these strains carry no other mitochondrial DNA) and hybridization experiments show a sequence homology of only 200 to 250 base pairs (bp) between the 1,900 to 2,000 bp var1 transcript and this petite mitochondrial DNA. These observations have been used to suggest, by elimination, that the *VAR1* gene is located in the nucleus; a directly testable hypothesis.

The search for the promoter region occupied the attention of those trying to unravel the control of primary transcription. The results emphasized once again that control of gene expression in eukaryotes differs markedly from that in prokaryotes. A deletion ending 140 bp upstream from the ATG of the *his3* gene (transcription begins at position −45) resulted in loss of gene function as did a deletion ending 250 bp upstream from the ATG of the *cyc1* gene. Insertion of a *Ty1* sequence 161 bp upstream from the ATG of the *his4* gene (transcription begins at position −63) similarly eliminated transcription. In the case of the *gal* gene cluster, transcription of the contiguous *gal* 1 and *gal 10* genes diverge from a single region of approximately 700 bp. It is not yet known whether or not transcription of the two genes is initiated from the same point. From these results it appears that a surprisingly large region might be required for effective promotion of transcription. However, it is important to note that deletions do not simply remove stretches of DNA. They also fuse the flanking sequences that remain and it may be important to ascertain the consequences of these fusions.

In a related area, four laboratories reported the structural characteristics of cloned genes (*adh2, car1, cyc2, dur1,2*) derived from mutant strains in which production of an enzyme was rendered constitutive by a mutation situated close to the structural gene. Constitutivity at these loci was observed in *MATa* and *MATα* haploid strains and *MATa/MATa* and *MATα/MATα* diploid organisms, but constitutivity was lost in *MATa/MATα* diploids. In each case analysed there appeared to be an insertion in or near the mutated gene. In two instances the inserted DNA was shown to be related to *Ty* sequences. Pertinent to these observations is the fact that RNA hybridizing to *Ty* elements has been reported to be the most abundant transcript expressed in parallel with mating-type specific functions, that is, such RNA is present in *MATa* and *MATα* cells, but is diminished in *MATa/MATα* diploid strains. While this raises the possibility that the constitutively expressed genes have been placed under the control of a new promoter as a result of the DNA rearrangement, transcripts derived from wild-type and constitutive *ADH2* mutants exhibiting the mating-type effect have been compared on northern blots and found to be the same size. It is also important to point out that insertion of *Ty* sequences into the 'control' region of a gene does not always result in its constitutive expression or association of expression with mating-type functions. Two mutants unable to express the *his4* gene (*his4–912* and *his4–917*) were found to harbour *Ty* elements upstream from the transcribed sequences. A His$^-$ phenotype was observed both in haploid and homozygous *MATa/MATα* diploid versions of these mutants.

Association of *Ty* sequences with a major proportion of the chromosomal rearrangements analysed so far prompted a search for other genetic elements associated with the influence of these mobile elements on genes in their proximity. This was done by selecting His$^+$ revertants of the *his4–912* and *his4–917* mutants just discussed. One class of mutants (*spm2* and *spm3*) suppressed the effects of the *his4–917* mutation, but were incapable of suppressing the *his4–912* lesion. However, the reversion frequency of the latter mutation was increased 100-fold in Spm2$^-$ and Spm3$^-$ strains. Analysis of representative revertants revealed that recombination had occurred between the two directly repeated δ sequences flanking the inserted *Ty* element leaving behind a single copy of the δ sequence in place of the entire *Ty* element. Such revertants were found to be cold sensitive in a wild-type background, i.e. His$^+$ at 37°C, but His$^-$ at 22°C. In Spm2$^-$ or Spm3$^-$ strains cold sensitivity of the revertants was suppressed. A second class of mutations (*spm1*) suppressed the effects of both *his4–912* and *his4–917* mutations and cold sensitivity of the *his4–912* revertants. In formal terms these observations share several characteristics in common with the two-component *spm* control system of maize.

Several landmark events were identified for sporulation. Premeiotic DNA synthesis occupied the first 1.5–2 hours. Meiotic recombination was shown by direct measurement to occur at 4-6 hours, thus correlating with the pachytene stage. A novel and useful system for the manipulation of recombination-deficient mutants was presented. It has been difficult to diagnose recombination deficiency because of non-disjunction at meiosis I and the resulting spore inviability. These problems are circumvented with the *spo13* mutation. Strains harbouring such a mutation enter meiosis, undergo recombination and then, by-passing the reductional division of meiosis I, proceed

---

Terrance G. Cooper is an Associate Professor in the Department of Biological Sciences at the University of Pittsburgh.

0028-0836/81/020119-02$01.00

directly to a single equational division of meiosis II. The latter division does not require recombination. Therefore, by coupling *spo13* to a putative recombination-deficient mutation, it is possible to follow recombination and sporulation independently.

Gene conversion was suggested to be the basic signal of meiotic recombination. The different types of segregation observed are then a function of how the recombination-heteroduplex is corrected. The *rad52* gene product was shown to be required for many recombination-associated processes including: mating-type switching, meiotic recombination, γ-radiation repair, meiotic and mitotic gene conversion. However, sister-strand exchange, plasmid integration into a chromosome and excision of an integrated plasmid occur normally in Rad52⁻ strains. Mutants defective in the latter two processes were isolated and in one case found to be defective in all forms of mitotic recombination assayed. Transformation requiring plasmid integration (Ars⁻) was 25-fold lower in this strain. Also isolated were four classes of mutants (*cor*) exhibiting elevated frequencies of meiotic recombination. Cor⁻ mutants were not sensitive to radiation and did not affect meiotic map distances or the frequency of initiating gene conversion. However, correction of heteroduplex intermediates generated during recombination was markedly reduced and the length of the correction tracts appreciably reduced.

The phenomenology and genetics of DNA repair is rather complex. Approximately 100 mutants have been isolated and characterized. This, however, may represent an overestimate of the genetic elements participating in repair, because allelism tests have not been completed. The results of characterization studies in many laboratories have led to the conclusion that five repair systems seem to exist. One mutagenic (error-prone) and two non-mutagenic systems are involved in the repair of radiation damage. Work with anti-mutator mutants suggests the existence of two mutagenic repair systems. One of these systems is identical with the mutagenic repair system associated with radiation damage. However, the second system appears to be independent of radiation damage. A fifth system appears to be needed for the repair of damage generated by the DNA alkylating agent, methyl methane sulphonate (MMS). This suggestion is based on the observation that some MMS-sensitive mutants are insensitive to radiation. The repair of psoralen-generated cross-linked DNA requires both non-mutagenic repair systems associated with repair of radiation lesions (the *rad3* and *rad51* systems). A new assay system has also been developed which permits one to follow the repair of DNA cross-links in 2 μm circular DNA. This system should prove useful for studying DNA repair at the molecular level.

Finally, two reports describing very interesting types of DNA were presented. First was the centromeric DNA (*CEN3*) from chromosome III which was isolated on a 1.6 kb segment of DNA located near the centromere-linked *cdc10* locus. When present on a plasmid carrying a yeast chromosomal replicator (such as *ars1*), *CEN3* DNA enabled the plasmid to segregate as a chromosome both mitotically and meiotically. In meiosis the plasmid segrated 2+:2− rather than the 4+:0− characteristic of other unintegrated plasmids. In addition, plasmids containing centromeric DNA were found to be strikingly more stable. This stability would be a most desirable characteristic in transformation cloning vectors. Such vectors are now being constructed. The second interesting type of DNA was reported to exist as a naturally occurring plasmid in *Kluyveromyces lactis*. This organism appears to harbour two linear DNA plasmids of molecular weights 5.4 and 8.4 × 10⁶ respectively. They represent the first non-circular DNA plasmids reported in yeast or any other organism. Strains possessing the two linear plasmids were shown to kill both *S. cerevisiae* killer and sensitive strains. Non-killing *K. lactis* strains both sensitive and resistant to being killed themselves were also found.

Remote sensing of the oceans

from D.J. Webb

NEW techniques for remote sensing of the ocean surface have been tested in two cooperative experiments — the Marineland experiment off the coast of Florida and the West coast experiment off California — and the results reported in a recent issue of the *Journal of Geophysical Reseach* (**85**, 1980). The experiments were carried out in 1975 and 1977 and used centimetre radar, high frequency (HF) radar, optical and surface buoy measurements.

A variety of techniques have been successful in sensing properties of the oceans. Sound can be used at long range — sonar techniques give bottom topography, doppler shifts give the water velocity and changes in sound paths detect the movement of density surfaces. The use of electromagnetic sensors is limited by the high conductivity of sea water but comes into its own for monitoring the sea surface. Passive infra-red sensors can monitor sea surface temperatures and optical techniques can be used to measure the slope of ocean waves.

Recently, active sytems have been developed, many using radar waves of between three and thirty cm wavelength. When directed vertically onto the sea surface the radar waves are scattered in a manner analogous to Sun glitter. In the SEASAT-1 radar altimeter the scatter can be used to give information on wave height and large-scale undulations of the sea surface[1]. If altimeter and positional errors can be reduced below 10 cm such a system can also be used to detect tides and strong surface currents.

When radar is used at an angle of more than 30° from the vertical, reflection occurs mainly by Bragg scattering from the short surface waves. The height of these waves and thus the back-scattered cross-section depends on the wind field, the long waves they are riding on and possibly on the turbulence and surface tension of the surrounding water. Any changes in these underlying parameters may be detected by the radar signal.

HF radar using wavelengths of tens to hundreds of metres may also be used. The radar signal is Bragg scattered by ocean waves of half the radar wavelength and interpretation of the returned signal is relatively simple. The method may be used to measure both the directional wave field and, from the doppler shifts, the surface currents, over distances of tens to hundreds of kilometres from the transmitter.

However, HF radar cannot be used from satellites because the signal is reflected by the ionosphere. Centimetre radars can be used but it is necessary to make sense of the complicated physics involved in the scattering process. It was this problem that led to the setting up of the 1975 and 1977 co-operative experiments.

In the first experiment[2], it was found that during a period of wave growth the power spectra for waves of between 1 and 5 seconds period did not depend on frequency to the power of −5, as the text books say, but instead the exponent lay between −3.5 and −4.5. The same experimental group using HF radar, with its accurate directional resolution, found evidence for the Phillips' mechanism of wave growth affecting waves of up to 25 m wavelength. These are rather larger than the catspaws on still water to which the theory is usually applied.

The HF technique was also used to study wave sheltering by islands[3]. A local Loran-A transmitter with a wavelength of 154 m was used and a directional aerial synthesized by driving a van at a fixed speed along nearly straight stretches of a suitable highway. The study showed a reduction in the wave energy behind the islands, but the

---

D.J. Webb is with the NERC Institute of Oceanographic Sciences, Godalming, UK.

0028-0836/81/020120-02$01.00

residual energy was greater than could be explained by refraction or diffraction. The authors suggest that non-linear wave interactions may be responsible.

The island study also made use of synthetic aperture radar (SAR) using centimetre waves. A similar system used on SEASAT-1 had an all-weather resolution of 20 m, and around the UK showed swell waves off Northern Scotland, ships and their wakes in the Straits of Dover, disturbances produced by sand waves in the North Sea and internal tides propagating out into the Atlantic from the continental shelf edge.

The classical explanation of why swell waves are seen by such a system is that their orbital motion strains the short gravity waves and changes their amplitude. However, Wright and his co-workers[4] found this explanation to be incorrect for typical ocean waves and suggest that it is the local wind field around the waves which is responsible for the observed signal.

Pawka and co-workers[5] compared the directional wave spectrum obtained from SAR images with the spectra from a high resolution array of pressure sensors. They concluded that when waves could be imaged, the SAR gave useful information on the direction and directional spread of the waves.

Similar results were obtained by McLeish *et al.*[6]. This group also found that, when normalized at the peak, the high frequency part of the SAR spectrum lay between the spectra obtained from wave height and orbital velocity measurements. The SAR system also showed swell being refracted by the Gulf Stream[7] and swell being refracted and scattered by bottom topography as it approached the shore[6].

When the same system was used in the much lower resolution scatterometer mode[8], some gross changes were found on crossing the Gulf Stream. The authors conclude that this was due to changes in the surface stress due to the wind. This could arise because of the lower atmospheric stability over the Gulf Stream.

Finally, scatterometer measurements over a larger area were used to study how the back-scattered cross-section depends on the relative direction of the wind[9]. A much larger cross-section is observed when looking up or down wind than when looking across the wind, and a development of this technique was successfully used in the recent SEASAT satellite to measure the wind field over the ocean.

When the experiments were carried out the radar systems were all at an early stage of development. Since then the altimeter, scatterometer and SAR have all been successfully used in the SEASAT-1 satellite and commercial versions of the HF radar are being developed for coastal monitoring. Versions of these instruments are also planned for future satellites and one might predict that over the next 10 to 15 years the altimeter and scatterometer will become the standard instruments for measuring wave climate and winds over the oceans.

---

1. Tapley, B.D. *et al. Science* **204**, 1410 (1979).
2. Trizna, D.B. *et al. J. geophys. Res.* **85**, 4946 (1980).
3. Vesecky, J.F. *et al. J. geophys. Res.* **85**, 4977 (1980).
4. Wright, J.W. *et al. J. geophys. Res.* **85**, 4957 (1980).
5. Pawka, S.S. *et al. J. geophys. Res.* **85**, 4987 (1980).
6. McLeish, W. *et al. J. geophys. Res.***85**, 5003 (1980).
7. Hayes, J.G. *J. geophys. Res.* **85**(C9), 5025, (1980).
8. Weissman, D.E. *et al. J. geophys. Res.* **85**(C9), 5032 (1980).
9. Brucks, J.T. *et al. J. geophys. Res.* **85**(C9), 4967 (1980).

Multiple methylation and bacterial adaptation

from Judith Armitage

THE ability of bacteria to sense, respond to and adapt to chemical stimuli provides a simple model system for the study of sensory transduction is general. It has been known for some time that protein methylation is a key factor in the process of adaptation to new chemical environments and now, a series of papers shows that methylation reactions are more complicated than had been previously thought and allow a much more flexible range of responses.

In an unchanging environment bacteria swim in a random fashion with periods of smooth swimming interrupted by transient bouts of tumbling, caused by reversal of the direction of flagellar rotation. When a gradient of an attractant or a repellent is encountered, tumbling is suppressed when the bacteria are moving towards the attraction or away from the repellent resulting in overall migration towards positive or away from negative stimuli. A sudden change in attractant or repellent concentration produces similar behaviour; the small size of bacteria requires that they sense all gradients as temporal rather than spatial. However, in this case, after a short time bacteria adapt to the stimulus and resume swimming even though the chemical is still present.

If bacteria are starved of methionine they can still respond to changes in the environment but they lose this ability to adapt. Mutants with aberrant behaviour, either tumbling continually (*Escherichia coli, che* B) or never tumbling (*E. coli, che* X) have also been found. The mutants can, however, be induced either to swim smoothly, or to tumble but fail to then adapt; *che* X mutants will continue to tumble for hours after repellent stimulation. The mutants proved defective in enzymes required to methylate or demethylate a specific group of inner membrane proteins.

These proteins, the methyl-accepting chemotaxis proteins (MCPs), are now known to be responsible for adaptation. Methyl groups are transferred to glutamyl residues on the MCPs from s-adenosyl methionine, in the presence of attractants, using a specific methyl transferase, creating a glutamyl methyl ester, or are removed in the presence of repellents using a methyl esterase. A high level of methylation is maintained as long as the attractant is present at a constant concentration, blocking the signal from the bound receptors. There are certainly two and probably three different groups of MCPs in *E. coli* and *Salmonella typhimurium* with molecular weights between 50,000 and 65,000, each cell probably having several hundred copies of each. Each group of MCPs responds to a specific set of chemosensory receptors which may be part of or able to bind to the MCPs and they provide parallel, almost independent pathways from their respective groups of receptors to the flagella.

Initially it was assumed that there were just two forms of MCP, one methylated and active, the other demethylated and inactive, controlling the function of a so-called 'tumble-generator'. However, on high-resolution polyacrylamide gels each MCP runs as a characteristic set of multiple bands.

While devising a method for methylating MCPs directly with labelled methionine in permeabilized cells, Paoni and Koshland (*Proc. natn. Acad. Sci. U.S.A.* **76**, 369; 1979) noticed that addition of attractants not only increased the intensity of MCP labelling on SDS gels but that the pattern of labelling within the bands changed, suggesting sequential methylation.

This observation has now been followed up in three papers published almost simultaneously and producing what seems to be at least part of the answer to the multiple banding on SDS gels and to adaptation, but showing that adaptation in bacteria is not the simple biochemical switch previously thought.

All three papers show that the multiple banding pattern on SDS gels is the result of sequential methylation of specific sites on the MCP polypeptide. Engström and Hazelbauer (*Cell* **20**, 165; 1980) used *in vivo* labelling and two-dimensional gel

Judith Armitage is in the Department of Botany and Microbiology, University College London.

0028-0836/81/020121-02$01.00

electrophoresis. The other two groups, DeFranco and Koshland (*Proc. natn. Acad. Sci. U.S.A.* 77, 2429; 1980) and Chelsky and Dahlquist (*Proc. natn. Acad. Sci. U.S.A.* 77, 2434; 1980) used recombinant genetics to introduce phage-directed cloned MCP-coding genes into *E.coli* and examined the labelling pattern of these MCPs in permeabilized cells. All the papers showed that stimulation with an attractant resulted in more heavily labelled faster-migrating bands, often with the appearance of a new fast band. However digestion of each band gave the same peptide pattern. At least three distinct peptides were shown to be methylated. Complete backbone labelling of the MCP by DeFranco and Koshland also revealed an additional slower-migrating band, corresponding to the unmethylated protein.

Sequential methylation of the MCP in response to an attractant, or demethylation in response to a repellent suggests that the number of sites methylated corresponds in some way to the number of receptor sites bound. The MCPs remain at the same level of methylation as long as a specific number of receptor sites is filled, inhibiting further response to that chemical at that concentration, but allowing responses to changing levels of that chemical or to new stimuli up to saturation. This obviously allows flexibility within an otherwise limited system. Each MCP must, however, be able to interact with more than one receptor molecule at a time, or the stoichiometry of the receptor per MCP methylation must be greater than one.

Interestingly, both Hazelbauer and Dalquist showed that chemoattractants such as maltose and aspartate which produce different levels of bacterial response also gave different MCP banding patterns. This may help to explain why attractants with similar numbers of receptor proteins vary in their efficiency in eliciting a response.

The question remains of how the binding of a chemical to the receptor–MCP complex controls sequential methylation or demethylation. A recent paper from Adler's laboratory may provide the answer (Black, Hobson & Adler *Proc. natn. Acad. Sci. U.S.A.* 77, 3879; 1980). They show that fluctuations in internal levels of cyclic GMP may initiate adaptation. Mutants showing defective adaptation also have defective cyclic GMP metabolism. Also, *in vitro* demethylation was inhibited by addition of cyclic GMP and cyclic GMP caused smoth swimming in the absence of chemoeffectors.

As yet the transducing signal between MCPs and flagella is unknown. It is almost certainly another function of the MCP but whether it is the release of a small molecule, possibly cyclic GMP as suggested by Adler, induced by a conformational protein change on binding, or a transient change in membrane potential or charge is not clear. It will also be interesting to discover whether bacteria are able to adapt to other tactic stimuli. Covalent modification of proteins has been found in several sensory systems in higher organisms, rhodopsin for example undergoes multiple phosphorylation when photostimulated and the methyl transferase system is involved in both stimulation of hormonal secretions in some higher organisms and leukocyte taxis. Sequential modification of proteins may therefore be a general biological phenomenon in signal tranduction.

Photosynthesis in the Greek sunshine

from B. Halliwell and M. C. W. Evans

THE rapid advance in our understanding of the photosynthetic reaction centres has come about mainly through work on bacterial systems but knowledge of the plant reaction centres is now fast catching up. Components analogous to those in bacteria have been identified in photosystem II. Ke and Klimow (Yellow Springs, Ohio) have detected a pheophytin intermediary electron carrier, which was also reported in barley mutants by Nugent and co-workers (University College London). There is strong evidence that the subsequent acceptor is a quinone probably associated with an iron atom. Manipulation of the redox states of these components gives rise to EPR signals, absorption spectra and chlorophyll triplet states, indicating a mechanism much like that in purple bacteria (Rutherford, Urbana). Nugent *et al.*'s work on barley mutants also showed the importance of the donor to P680 as a source of the EPR signal of Photosystem II.

In Photosystem I, reviewed by Ke, EPR and endor measurements have shown that the pheophytin intermediate is replaced by a chlorophyll *a* monomer and the quinone acceptors by a group of iron-sulphur centres. Understanding of the purple bacteria is much more advanced, however. For example, Shuvalov (Puschino, USSR) and Parson (Washington) proposed a role for the B800 chlorophylls in the sub-picosecond primary photochemistry. Our knowledge of electron transfer outside the reaction centre is also best in the purple bacteria. A specific quinone involved in the oxidation of the Rieske iron-sulphur centre has been identified. This seems to exclude a Q-cycle system and requires a linear electron transport chain. However, a paired electron flow is indicated by the gating effect which can be observed in the oxidation and reduction of the secondary quinone acceptor (Q_2) and how these observations can be reconciled is not yet clear.

In chloroplasts the isolation of a cytochrome *bf* complex, analagous to the *bc* complex of mitochondria, suggests a similar mechanism in the quinone oxidation system. Since Z remains unidentified, however, the problem of a Q cycle versus a linear electron chain is unresolved. Different schemes were proposed by Bouges-Boucquet (Paris), Rich and Bendall (Cambridge, England) and Hind (Brookhaven). The problem is further complicted by the conflicting requirements of cyclic and non-cyclic electron transport in chloroplasts. As discussed by Bendall, there is no definitive explanation of the function of the *b* cytochromes, particularly the two forms of b_{559}. Hence, while most participants at the Congress continued to accept the 'Z-scheme' of two light reactions in series, alternative arrangements have been proposed (for example, by Arnon at Berkeley, California). Indeed, there is room for some modification, since according to Bjorkman (Stanford) the relative numbers of photosystem I and II reaction centres and electron transport chains can vary with the growth conditions of the plant.

Perhaps the most disappointing part of the Congress was that dealing with water oxidation, which is still the least-well

---

*The Fifth International Congress on Photosynthesis was held at Halkidiki, (near Thessaloniki), Greece, from September 7th to 12th 1980, in appropriately sunny weather. The Congress dealt with all aspects of research in photosynthesis, from physiological ecology through carbon metabolism to the extreme limits of time resolution in the sub-picosecond photochemistry of the photosynthetic reaction centres.

M. C. W. Evans is in the Department of Botany and Microbiology at University College London and B. Halliwell is in the Department of Biochemistry at King's College London.

0028-0836/81/020122-03$01.00

Interstellar media in other galaxies

from a correspondent

A recent international workshop at the European Space Agency's (ESA) Villafranca tracking station (Vilspa) near Madrid brought together those studying the interstellar medium in our own galaxy and those concerned with extragalactic problems, in particular the quasar absorbtion lines. Data from both the International Ultraviolet Explorer (*IUE*) satellite, which is controlled from Vilspa during its European observing shift, and Copernicus, as well as from ground based telescopes were discussed. The *IUE* has just begun making important observation of interstellar media and, with the demise of Copernicus at the end of 1980, will be the prime instrument for interstellar studies until the launch of Space Telescope.

Ted Snow (Colorado) reviewed work on the abundances of elements and their depletion in the interstellar gas as a result of intensive studies of particular sight-lines with the Copernicus ultraviolet spectrometer. While gas-phase abundances of some species, in particular carbon and oxygen, are little depleted and this sets limits on grain composition, the refractory elements are more highly depleted, consistent with either grain condensation in a high-density environment or preferential accretion of these species onto grain mantles by dynamic processes in interstellar clouds. Stuart Pottasch (Groningen) noted that, for a few species, the existence or degree of depletion depended on the cosmic abundances adopted and these were themselves not beyond controversy. Nolan Walborn (Cerro Tololo) reported some recent work on the dynamics of the interstellar gas in the η Carina nebula; relative velocities up to 550 km s^{-1} are seen. Ed Jenkins (Princeton) discussed the presence of highly ionized gas in the plane of the galaxy. It appears that OVI lines define a component with scale height above the galactic plane of $\sim$300 pc and a velocity dispersion of 25 km s^{-1}. Their strength is correlated with reddening by dust. By contrast there is still some controversy whether the CIV and SiIV lines are truly interstellar. There seemed some evidence that the latter, at least, was in part circumstellar. Gordon Bromage (Chilton) claimed that NV lines were not seen in *IUE* data covering sight-lines to halo stars which rise out of the galactic plane so that the NV distribution looks similar to that of OVI with a scale height much smaller than that of the lower temperature CIV near 3kpc. Max Pettini (Herstmonceux) noted, however, that he had one case where NV was seen in a halo star and he added that the three-time ionized carbon abundance may be lower in the plane than just outside it.

Klaas de Boer (Washburn) reviewed the recent exciting work by *IUE* on the galactic halo with emphasis on sight-lines to the Magellanic Clouds. There is some uncertainty if gas far from the plane co-rotates with the plane but, if so, the density falls by a factor of 10 over 10 kpc and one cloud covering 7 square degrees in the direction of the Large Cloud may contain 300,000 solar masses. There is component structure in OI and FeII but the SiIV and, possibly, CIV profiles suggested the halo has a smoother, less cloudy structure than the plane. There was, however, some discussion as to whether the direction to the Magellanic Clouds (where most of the *IUE* data exists) is typical of our galaxy. Claudine Laurent (Verrieres-le-Buisson) commented also that the existence of a similar halo around the Small Cloud was in doubt as she had data showing no CIV or SiIV in an SMC star which, unlike other cases observed to date, did not have its own HII region. Chris Blades (Chilton) reported ground-based optical data from the Anglo-Australian Telescope (AAT) and found that the ratio of the neutral sodium column to dust reddening in the Small Cloud was 5 times smaller than that in our own galaxy.

Michael Penston (Vilspa) also showed AAT data of optical interstellar lines in other galaxies using galactic nuclei or supernovae as background probes. He confirmed earlier reports of CaII absorption in the Magellanic Stream. In discussion, the point was made that the (by necessity) moderate resolution on optical interstellar lines in other galaxies is similar in quality to the type of data that was available for our own galaxy in the 1940s and 1950s. While it is valid to analyse the galactic and extragalactic doublet ratios is similar ways and show that turbulent motions in interstellar media are generally characterised by velocity dispersions similar to that in our own galaxy, it was agreed that it is dangerous to deduce column densities from such material because large columns can be 'hidden' in components with low velocity dispersion. Alec Boksenberg (London) reviewed the current status of work on quasar absorptions lines. The absorption systems with metal lines can arise in galactic halos if these have radii of 100 kpc. The 'Lyman-α-only' absorption systems are not clustered like galaxies and suggest the presence of intergalactic clouds of primordial material.

On the theoretical side, Jo Silk (Berkeley) discussed other less direct evidence for intergalactic clouds including deduction from theories of galaxy formation. He proposed that slow self-regulating star-formation takes place in these clouds prior to their merging into galaxies as we know them. Joel Bregman (New York) described his 'galactic fountain' model of galactic haloes, in which hot gas of temperature T_0 at the base of the galactic corona rises, cools and turns into clouds which fall back towards the plane. Coronae of the right size are produced if T_0 is in the range one to six million degrees. A total mass of 10^{10} solar masses is required in these galactic haloes.

---

understood part of the electron transport system. Recent reports from Winget (Cincinnati) on reconstitution of this system using an isolated manganese-protein were not substantiated by other workers. For example, the reconstituted system described by Nakatani *et al.* (Imperial College London) requires a catalase-like protein rather than a manganese protein. Also, the NMR measurements of Wyrdzinski and Govindjee (Urbana) that appeared to reflect changes in the redox state of manganese were challenged by Sharp and Yocum (Ann Arbor, Michigan). Nevertheless, materials for the solution of the 'water-splitting problem' are accumulating. Oxygen-evolving preparations with increased specific activity were described by Stewart (Cambridge, England), inverted vesicles, that is, with the water oxidation system on the outside, were described by Albertsson (Lund), and Bishop (Corvallis, Oregon) has discovered mutants with defective water-oxidation systems. We cannot yet be 100 per cent sure, however, according to Metzner (Tubingen), that it is H_2O and not CO_2 that is releasing oxygen!

The sessions on chloroplast metabolism

0028-0836/81/020123-01$01.00

dealt largely with the regulation of the Calvin cycle in C_3 plants, starch and sucrose synthesis, photorespiration and nitrogen metabolism. Few advances in our knowledge of C_4 plants were reported, although both Sugiyama (Shizuoka) and Hatch (Canberra) reported that regulation of the activation and deactivation of pyruvate-phosphate dikinase in the mesophyll cells is achieved by adenine nucleotides rather than thiol-dependent mechanisms. The pathway of chlorophyll biosynthesis in chloroplasts, particularly the route from glutamate to δ-aminolaevulinic acid, is becoming clearer. Chlorophylls *a* and *b* with modified side chains in the ring structure also occur in leaf tissues (Rebeiz, Urbana).

Regulation of the Calvin cycle is achieved by several mechanisms. One possibility proposed by Buchanan (Berkeley) and Anderson (Chicago) is a light-dependent increase in the activity of certain enzymes. In spinach chloroplasts, illumination generates reduced ferredoxin which produces, via a reductase enzyme, the reduced form of the thiol protein thioredoxin. Three types of thioredoxin (m_b, m_c and *f*) have been found in chloroplasts by Schurmann (Neuchatel). Reduced thioredoxin *f* can increase the activity of some Calvin cycle enzymes assayed at physiological substrate and cofactor concentrations, and rapid rises in the activity of fructose and sedoheptulose diphosphatases on illumination of chloroplasts were reported by Leegood (Sheffield) and by Heldt (Gottingen). According to Charles (King's College London) in the case of fructose diphosphatase the rise in activity is due to an increase in the affinity of the enzyme for its substrate and cofactor rather than to a rise in V_{max}. When the light is turned off enzyme activities fall. How this is achieved has not been established. Heldt and others have shown that illuminated chloroplasts generate H_2O_2, which deactivates the enzymes and may account for the lowered rate of CO_2 fixation in chloroplast preparations illuminated in the absence of added catalase. In the presence of catalase, however, the rises in enzyme activity on illumination are too fast to limit the rate of CO_2 fixation both before and after the induction period (Walker, Sheffield). Ribulose diphosphate carboxylase shows little light activation. Hence CO_2 fixation must be controlled by other means *in vivo*. This led Halliwell (King's College London) to propose that the thioredoxin system exists to protect the reduced, physiologically active forms of certain Calvin cycle enzymes from H_2O_2 generated *in vivo*, rather than as a regulatory mechanism. According to Heber (Dusseldorf), pseudocyclic photophosphorylation, generating H_2O_2, must occur *in vivo* to correctly poise the redox state of the electron transport chain.

The major factors regulating CO_2 fixation seem to be the rise in *p*H of the stroma on illumination, and possibly a light-dependent increase in the stromal Mg^{2+} ion concentration. Increases in stromal [Mg^{2+}] from 1 to 3 mM have been reported in spinach chloroplasts, as discussed by Heldt, but Ben-Hayyim (Bet Dagan, Israel) has found no such increase in lettuce chloroplasts. There were some speculations that chloroplast Ca^{2+} ions, which powerfully inhibit fructose and sedoheptulose diphosphatases and protein synthesis, might be involved in turning off these reactions in the dark (for example, Jagendorf, Cornell).

Further developments in our knowledge of starch degradation in chloroplasts were reported by Stitt (Gottingen), who has isolated starch-loaded chloroplasts capable of degrading their starch at a physiological rate. Both amylase and phosphorylase activities are involved in this degradation (Beck, Bayreuth). Foyer (Sheffield) reported studies on the purified forms of the enzymes of sucrose synthesis. Claims by Garrett (Belfast) that isoenzymes of ribulose diphosphate carboxylase from diploid and tetraploid strains of ryegrass have different affinities for CO_2 and O_2 were fiercely contested, and more work is needed on this system.

The glycine–serine pathway is still regarded as the major pathway of CO_2 release during photorespiration in leaves, a view that was strengthened by the description of mutants of *Arabidopsis* with defects in various enzymes of the pathway (Ogren, Illinois). The ammonia generated during glycine decarboxylation is, according to Miflin (Rothamsted), assimilated by glutamine synthetase in the leaf cytosol. Other pathways of photorespiratory CO_2 release cannot be completely ruled out, however, since CO_2 evolution by direct glyoxylate decarboxylation has been reported in isolated leaf cells (Oliver, Idaho and Grodzinski, Ontario) and has been suggested by Ogren to occur at a high rate in an *Arabidopsis* mutant which cannot provide amino acids for the transamination of glyoxylate to glycine. Perhaps these alternative pathways are only important in nitrogen-deficient leaves.

The Rothamsted group also presented evidence that chloroplasts can synthesize amino acids of the aspartate family (Thr, Lys, Ileu, Met). It is now well established that ammonia assimilation in chloroplasts proceeds by the glutamine synthetase/glutamate synthetase route. Glutamine produce by cytosolic glutamine synthetase (see above) must enter the chloroplast for further metabolism by the ferredoxin-dependent glutamate synthetase. □

The (p,α) reaction

from P.E. Hodgson

THERE have recently been several studies of nuclear reactions emitting particles in an unstable or resonant state. These particles live long enough to act as a single particle during the reaction itself, but break up into two fragments soon after emission and long before they reach the detecting equipment. For example, it is possible for ^{8}Be to be emitted from a nuclear reaction, but it is unstable and soon breaks up into two alpha-particles. The energy release is only 96 KeV, and so the alpha-particles continue to move in almost the same direction as the ^{8}Be nucleus and are easily identified by their strong directional and energy correlation compared with the single alpha-particles from other reactions. Another example of such a reaction is the (α, d*) reaction with the emitted deuteron in the unstable singlet state at an excitation energy of 100 KeV.

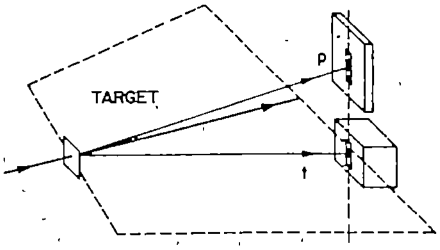


Fig. 1 Arrangement of detectors for the proton and triton from the breakup of excited alpha-particles.

It is particularly interesting to study such reactions as they extend the range of final states that can be reached and also provide an additional test of nuclear reaction theory. One reaction that was first identified a few years ago (*Nature* **266**, 121; 1977) is the (α,α*) reaction with the outgoing alpha-particle excited to the unbound monopole state at 20.1 MeV. Now a detailed study of the corresponding (p, α*) reaction has been made by Datta *et al.* at the University of Manitoba (*Phys. Rev.* **C21**, 1687, 1980). The excited alpha-particles decay into a proton and a triton, and these are detected by two counters as shown in Figure 1. The measurements were made with a 45.2 MeV proton beam on ^{12}C, and some of their results for the (p, α) and (p, α*) reactions are shown in Figure 2.

The cross-sections of many nuclear reactions can be calculated by the distorted wave theory, and this may be applied to

0028-0836/81/020124-02$01.00

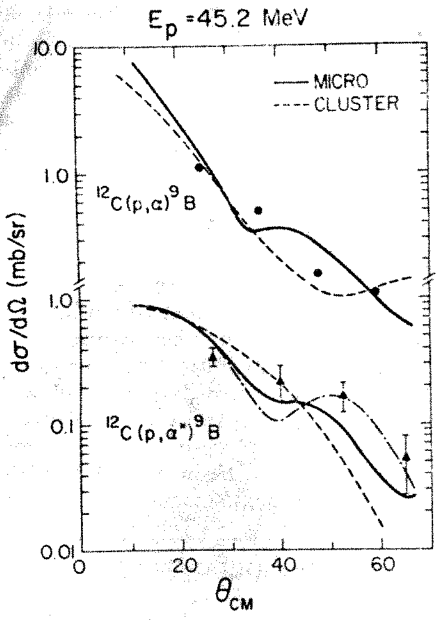


Fig. 2 Differential cross-sections for the (p,α) and (p,α*) reactions on ^{12}C at 45.2 MeV compared with distorted wave calculations with the microscopic and cluster models. Y-axis, differential cross section in mb/sr; X-axis, angle in degrees. The dash-dot curve in the lower figure shows the result of another microscopic calculation using different distorting potentials. Each calculation is separately normalized to the experimental data.

these reactions, provided allowance is made for the unbound nature of the emitted particle. This is done by expressing its wavefunction as a product of a bound s-wave function and an energy-dependent Breit-Wigner resonance factor. There are two possible ways of treating the three particles picked up by the incident proton from the target nucleus. The simplest is to treat them as a cluster, and then the theory appropriate to a one-nucleon transfer reaction can be used directly with appropriate modification for the mass and charge of the transferred cluster. The other method is to treat the transferred nucleons individually, which is more correct but considerably more complicated. The cross-sections corresponding to calculations made in both these ways are compared with the experimental data in Figure 2. It is notable that on the whole the calculations treating the nucleons individually, called the microscopic model, fit the data somewhat better, although there is little to choose between them. Both calculations reproduce quite well the general features of the cross-sections and also the ratio between the cross-sections of the *(p,α)* and (p,α*) reactions. There is some uncertainty connected with the choice of distorting potentials, but on the whole the agreement is sufficient to permit the conclusion that the (p,α*) reaction is sufficiently well understood for it to be used in studies of nuclear structure.

P.E. Hodgson is in the Department of Nuclear Physics, University of Oxford.

Precambrian orogenies and polar wandering

from J.C. Briden

[illegible] Library
Science College
Upper Circular Road

EVIDENCE that orogenic belts less than a few hundred million years old have formed (or are forming) at plate margins either at a continent/ocean interface or at a site of collision of two continents is overwhelming. It is consistent with the theory of a lithosphere made up of a small number of rigid plates in relative motion and encourages the interpretation of geology by extrapolation of the plate model far back in time. Extrapolation to more remote periods is uncertain especially because in Archaean times (beyond 3,000 Myr ago — the most remote parts of the geological record) evidence for rigid regions which might be the plate interiors is lacking. Hence notions have grown of intervals in which transitional tectonic regimes might have operated, with speculation that some or all older Precambrian orogenic belts do not mark former plate margins.

Palaeomagnetic studies have featured prominently in these arguments. The directions of permanent magnetism in continental rocks show slow systematic changes with age which can be depicted as paths of apparent wandering of the pole relative to each plate. For the past several hundred million years, the differences between the various polar wander paths are far more pronounced than any common component they may have, unequivocally demonstrating that the plates have been in motion relative to one another.

Returning to the Precambrian, much of the argument about the origin of orogenic belts relates to the degree to which polar wander paths from presently adjacent cratonic regions agree or disagree. The most serious problem is the impossibility of identifying the precise age of magnetization in Precambrian rocks, many of which have a complex history reflected in the frequent superposition of several palaeomagnetic components. Even where geochronological studies of the same rocks yield age estimates of various events in the rock history, uncertainties remain because events which leave palaeomagnetic imprint may not be synchronous with closure of isotopic systems in those rocks.

It has variously been argued (a) that the palaeomagnetic signatures from all fragments are essentially the same, being compatible either with true polar wander or coherent movement of a single super continent; (b) that they are all different and hence latter day plate tectonics was fully and exclusively operational as far back as the palaeomagnetic record goes, and (c) that similarities and differences both exist, and hence some orogenic belts mark sites of continental collision and others do not.

The supercontinent model requires a single common polar wander path. This path can only be regarded as internally consistent if many inexplicable errors in magnetic age assignment have been made. It is also extraordinarily long, even allowing for the fact that it is almost certainly already smoothed by the failure to resolve superposed magnetic components in many suites of rocks. It implies a Proterozoic history utterly different from that of more recent times — a world in which polar wander or drift of a single supercontinent operated instead of a relative plate motion; in which deformation within that supercontinent occurred instead of plate interiors being rigid; and yet a world in which motions of the lithosphere were considerably faster than in more recent time.

On the other hand it is much too easy to simply link the sequence of palaeomagnetic data from different cratons, note that they are different, and conclude that cratons behaved like latter-day plates. The polar wander paths will inevitably look different because in almost every case each path will have been sampled inadequately and at quite different instants of time.

The middle-of-the-road model presently seems to find the most favour with palaeomagnetists because some of the very few comparisons of well determined polar wander paths from neighbouring cratons really do seem to agree, for period of times which post-date the deformation of the orogenic belts now lying between them. However, this is scarcely worthy of being called a model at all, because it fails to tell us which orogenic belts are merely compressions within pre-existing plates and which had a pre-history of large relative motion of their margins. It does not predict the style of motion when displacement has been large; nor can it rule out small openings and closings where they have been small.

In this issue of *Nature*, Onstott and Hargraves (page 131) argue for a fourth style of large motion (thus exhausting the alternatives and ensuring that at least someone is right!) with adjacent cratons sliding past each other along transcurrent

---

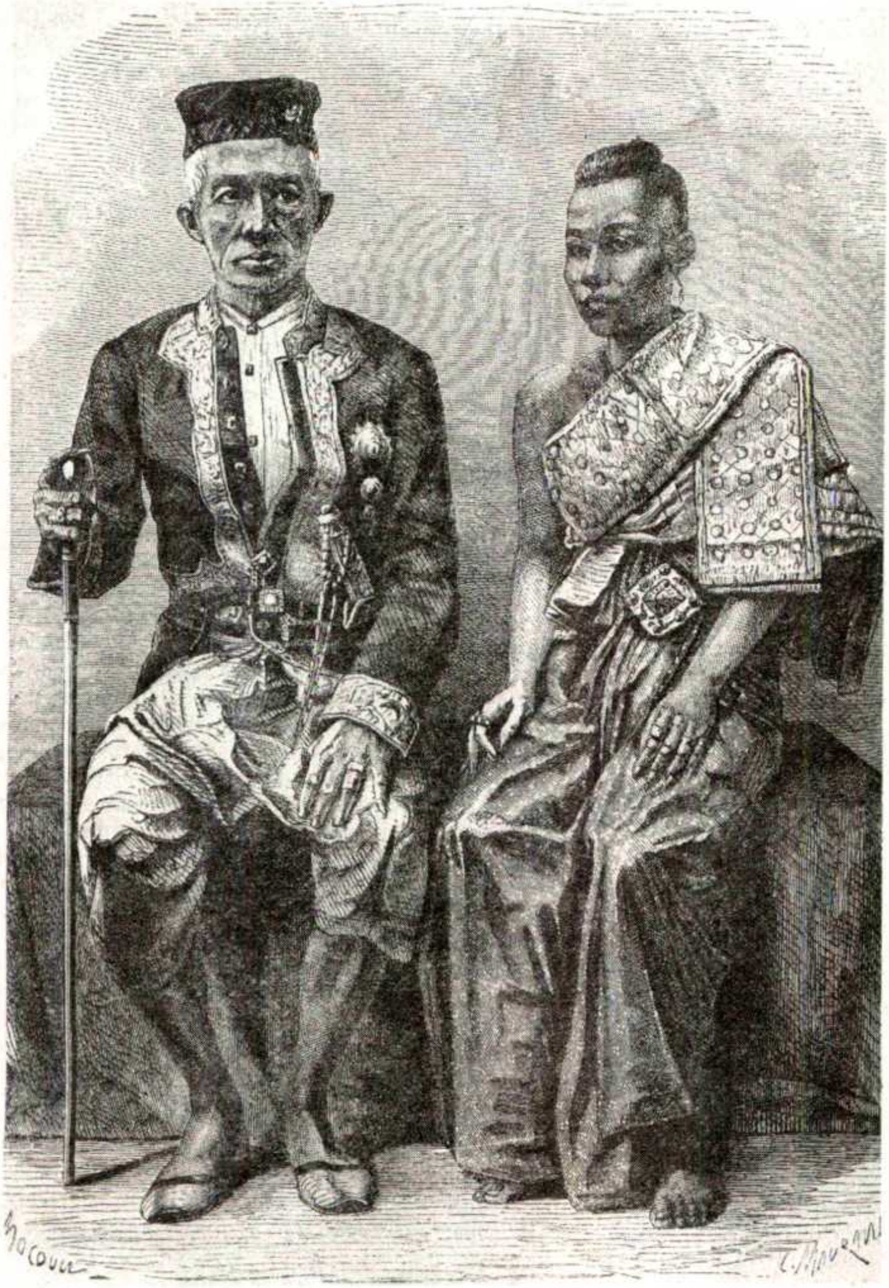
J.C. Briden is Professor in the Department of Earth Sciences, Leeds University.

0028-0836/81/020125-02$01.00

fault belts with neither creation nor consumption of oceanic crust within those belts — though presumably it has to happen elsewhere in other parts of the plate-driving system. Parts of their argument are as weak as any that have gone before due to the inadequacies and paucity of the palaeomagnetic data on which they have had to rely. Nor is their interpretation unique: their data are equally compatible with a 'Wilson cycle' of an ocean which opened and subsequently closed. Yet their model serves as an important reminder that this transcurrent style of tectonics is possible and draws attention to evidence consistent with it in Proterozoic times — there are substantial examples in the Phanerozoic, particularly in the evolution of the Caledonian and the Hercynian belts in Europe.

This element of relative plate motion is easily missed in tectonic analyses of orogenic belts as it is dependent on discovering a few surface traces of large faults parallel not only to the structural grain but most likely parallel to the grain of sedimentary facies as well. If one looks back at the palaeomagnetic evidence of different polar wander paths from Europe and North America published in the mid 1950s as evidence of the opening of the Atlantic, one would nowadays pick the individual data to shreds; few if any would be accepted as adequately confirmed palaeomagnetic directions by current criteria. Yet the correct answer could be seen, and was indeed seen, in the systematic deviation of sets of (internally suspect) data. The same may be so for Onstott and Hargraves' contention of large-scale transcurrent motions. As in the Phanerozoic the confirmation came from elsewhere than continental palaeomagnetism (in dated magnetic anomalies, transform fault interpretation, and the whole range of sea floor spreading data), so for the Proterozoic other disciplines may resolve the problem. The critical lines of inquiry are likely to be careful tectonic analysis of structures in plan, and seismic 'profiling' right through the lithosphere to give structure in cross-section. Mesozoic to present geological experience tells us that there are enough distinctions between major transforms and sites of destruction of major oceans or back arc basins for solutions to this problem to be attainable.

100 years ago



Mongoloid types, Indo-China. King and Queen of Siam.

MONGOLIAN TYPE

from A. H. Keane

We are all familiar with the essential characteristics of the Yellow or Mongolian division of the human family.

They correspond substantially with those we see in the ordinary Malay type, in Java, Bali, Madura, many parts of Sumatra, round the coast of Borneo, and in the peninsula of Malacca. The true aborigines of this region were the Negritos; consequently the Malays, like the pre-Malays or Caucasian Indonesians, are here intruders. Intruders from where? Obviously from where the type exists, the neighbouring Indo-Chinese peninsula. What then becomes of the Malay as a primary division of mankind? As such it can no longer be recognised in anthropology, and must sink to the position of a mere variety of the Mongol type. The so-called true Malay or typical Malay is essentially a Mongolian, and the likeness between the two has not failed to strike all careful observers. "The Malayan race," says Wallace, "as a whole undoubtedly very closely resembles the East Asian populations from Siam to Manchuria. I was much struck with this, when in the Island of Bali I saw Chinese traders who had adopted the costume of that country, and who could then hardly be distinguished from Malays; and on the other hand I have seen natives of Java who, as far as physiognomy was concerned, would pass very well for Chinese." Hence De Quatrefages rightly rejects the claim of the Malays to be regarded as a fundamental type. "All polygenists," he remarks, "have regarded the Malays as one of their *human species;* many monogenists have considered them as one of the principal races. I showed long ago that in reality they are only a mixed race in which white, black, and yellow elements are associated."

The last clause of this sentence gives the true solution of the problem. The inhabitants of Malaysia consist not of one, nor even of three distinct races, but of three races variously intermingled, the yellow or Mongolian, and the white or Caucasian chiefly in the west, these two and the black or Papûan chiefly in the east. As the fusion of yellow, white, and black produces the so-called "Alfuros" in the east, so the fusion of yellow and white produces the so-called Malays in the west. The more the yellow prevails the nearer do the Malays approach the Mongol type; the more the white prevails the nearer do they approach the Causasian type, until in some places they seem to be no longer distinguishable from the Mongols, in others from the Caucasians."

From *Nature* **23**, 13 January, 250, 1881.

0028-0836/81/020126-01$01.00

Fallacies of lifestyle cancer theories

Samuel S. Epstein* **& Joel B. Swartz†**

* Professor of Occupational and Environmental Medicine
† Assistant Professor of Occupational and Environmental Medicine,
School of Public Health, University of Illinois at the Medical Center, Chicago, Illinois 60680

Self-interest of the chemical industry apart, the lifestyle theory of cancer causation is held to reflect misinformation and 'inactive conservatism'.

PETO'S article[1], purporting to be a review of the book *The Politics of Cancer*[2], is largely a restatement of the lifestyle theory of cancer causation. This theory postulates that if you get cancer it is essentially your own fault, and that the causal role of past involuntary exposure to environmental and occupational carcinogens is trivial. Not surprisingly, the lifestyle theory has emerged as the major professed basis of the chemical industry's objections to the regulation of its carcinogenic products and processes[3]. As an enthusiastic proponent of this theory, Peto asserts that smoking-derived and fat-associated cancers "collectively account for more than half of all cancer deaths." As a corollary, of Peto's emphasis on lifestyle factors, he denigrates the role of occupational and environmental carcinogens and the need for their effective regulation, claiming that there has been no recent increase in cancer mortality rates other than that due to smoking. We shall demonstrate that there is scant scientific basis for the lifestyle theory, and that it is in fact contradicted by a substantial body of published evidence.

Cancer rate trends

Peto justifies his emphasis on lifestyle factors by dismissing evidence for recently increasing cancer rates, apart from "that due to the massive effects of smoking on lung cancer". However, there is substantial evidence to the contrary. Standardized cancer death rates, adjusted to the 1940 age structure of the total United States population, show a progressive overall increase of about 7% from 1935 to 1970 (ref. 4) despite marked reductions of stomach cancer rates for unexplained reasons and of cervix cancer rates for reasons including the frequency of elective hysterectomy for non-malignant disease and the success of screening programmes. These trends are consistent with standardized mortality data for the United States (Table 1)[5], where they are even more marked in black males, and with crude mortality data for the United Kingdom (Table 2)[6]. The overall rate of increase in US cancer mortality in the 7-year period from 1969 to 1976 (5.5%), adjusted to the 1970 age structure, is substantial and comparable with that for the preceding 35 years, 1935 to 1970 (7%). The overall increase in incidence rates is even more marked than mortality rates in the past decade, involving a wide range of organs besides the lung (Table 3)[5]. Moreover, the increase in incidence for all sites is comparable with that when lung cancer is excluded (Table 4)[7,8].

Reliance on overall age-adjusted incidence or mortality rates alone is simplistic, as such rates can mask steep increases in organ-specific cancers in high risk population subgroups, such as asbestos insulation workers or menopausal women treated with oestrogen replacement therapy. The overall probability, at today's death rates, of a person born now getting cancer by the age of 85 is 27% for both men and women; this is increased from the 19% for men and 22% for women born in 1950 (ref. 9). Furthermore, recent cancer rate trends reflect exposures and events beginning some 20 or 30 years ago, when the production of synthetic organic chemicals was relatively trivial compared with the present levels. The production of synthetic organic compounds in the United States in 1935, 1950 and 1975 was about 1, 30 and 300 billion pounds per annum, respectively[10]; sharp increases have also been observed for a wide range of derived industrial products such as chlorinated hydrocarbon solvents, plastics and resin materials, and of industrial carcinogens, such as vinyl chloride and acrylonitrile. It is reasonable to anticipate that greater production has been paralleled by increased exposure of increasing numbers of both the workforce and the general public, which is likely further to accentuate increasing trends in cancer rates. It must also be recognized that before the 1976 Toxic Substances Act, which the chemical industry so effectively stalled for so long[11], there were no requirements for testing chemicals before their introduction into commerce (with the exception of special-purpose legislation for drugs, pesticides and food additives). Thus, the overwhelming majority of industrial chemicals now in use have never been tested for chronic toxic and carcinogenic effects, let alone for ecological effects.

Role of smoking

As emphasized in *The Politics of Cancer*[1] (p. 178), "Smoking is the single most important cause of lung cancer, as well as of cancer at other sites, chronic bronchitis and emphysema, and cardiovascular diseases". Less well appreciated by lifestyle advocates is that overemphasis on smoking is widely used to divert attention from occupational causes of lung and other cancers. Of the approximately 100,000 annual lung cancer deaths in the United States, at least 20% occur in nonsmokers. It is relevant that lung cancer death rates in nonsmokers approximately doubled[12] from 1958 to 1969, an increase maintained since. Furthermore, the role of occupational exposure to carcinogens was not recognized in most of the classic epidemiological studies which linked lung cancer with smoking. This led to overestimation of the contribution of smoking compared with occupational risks or to their possible interactions.

Table 1 Age-adjusted cancer mortality rates per 100,000 US population for selected sites by sex and year 1969–76, and average per cent change[5]

		Mortality rate per 100,000		Average % change 1969–76	
Site	Sex*	1969	1976	Annual	7-Year
All sites	WM	195.0	210.2	0.9	7.8
	WF	129.0	133.8	0.5	3.7
Stomach	WM	10.6	8.7	−2.9	−17.4
	WF	5.3	4.1	−3.6	−22.6
Colon	WM	18.7	20.7	1.3	10.7
	WF	16.2	16.5	0.0	1.9
Rectum	WM	6.9	5.6	−3.0	−18.8
	WF	3.9	3.2	−3.1	−17.9
Pancreas	WM	11.0	11.0	0.0	0.0
	WF	6.6	6.8	0.3	3.0
Lung	WM	55.0	66.7	2.6	21.3
	WF	10.2	17.8	7.6	74.5
Melanoma	WM	2.0	2.6	4.0	30.0
	WF	1.4	1.6	0.8	14.3
Breast	WF	26.2	27.2	0.3	3.8
Cervix	WF	5.5	3.9	−4.9	−29.1
Uterus	WF	4.6	4.2	−1.7	10.5
Prostate	WM	19.0	21.0	1.2	8.7
Bladder	WM	7.1	7.5	0.6	5.6
	WF	2.1	2.0	−1.4	−4.8
Kidney	WM	4.3	4.5	0.6	4.7
	WF	2.0	2.1	0.7	5.0
Leukaemia	WM	9.4	9.2	−0.4	−2.1
	WF	5.7	5.2	−1.7	−8.8

For age adjustment the 1970 United States population was used as standard.
* WM, white male, WF, white female

0028-0836/81/030127—04$01.00

Table 2 Crude cancer mortality rates per 100,000 population for selected sites by sex and year, England and Wales, 1971–77, and average per cent change[6]

		Mortality rate per 100,000		Average % change 1971–77	
Site	Sex	1971	1977	Annual	6-Year
All sites	M	265	283	1.1	6.8
	F	215	233	1.3	8.4
Stomach	M	30	28	−1.2	−6.7
	F	21	19	−1.7	−9.5
Large intestine and rectum	M	30	32	1.1	6.7
	F	34	35	0.4	2.9
Respiratory	M	106	112	0.9	5.7
	F	22	29	5.3	31.8
Breast	F	45	47	0.7	4.4
Uterus	F	15	16	1.1	6.7
Prostate	M	17	19	1.9	11.8
Bladder	M	12	12	0.0	0.0
	F	4.1	5.1	3.7	24.4
Leukaemia	M	7.0	7.2	0.4	2.9
	F	5.5	6.0	1.5	9.1

Table 3 Age-adjusted cancer incidence rates per 100,000 US population (whites) for selected sites by sex and year, 1969–76, and average per cent change[5]

		Incidence rate per 100,000		Average % change 1969–76	
Site	Sex	1969	1976	Annual	7-Year
All sites	M	346.6	374.0	1.3	7.9
	F	271.5	301.2	2.0	10.9
Stomach	M	15.4	12.6	−2.3	−18.2
	F	7.1	5.6	−3.7	−21.1
Colon	M	34.5	36.9	1.5	7.0
	F	30.6	31.4	0.7	2.6
Rectum	M	17.5	19.4	1.3	10.9
	F	11.1	11.4	1.2	2.7
Pancreas	M	12.1	11.5	−0.5	−5.0
	F	7.5	8.0	0.9	6.7
Lung	M	70.6	77.8	1.4	10.2
	F	13.3	23.7	8.6	78.2
Melanoma	M	4.4	6.8	6.8	54.5
	F	4.1	6.1	6.2	48.8
Breast	F	73.9	83.5	1.8	13.0
Cervix	F	16.0	10.6	−5.9	−33.8
Uterus	F	22.6	31.2	5.9	38.1
Ovary	F	14.9	13.6	−0.4	8.7
Prostate	M	59.0	68.6	2.3	16.3
Bladder	M	23.8	26.4	2.3	10.9
	F	6.3	7.3	2.5	15.9
Kidney	M	9.0	9.6	1.2	6.7
	F	4.3	4.8	1.3	11.6
Leukaemia	M	13.2	13.1	−0.2	−0.8
	F	8.0	7.1	−1.0	−11.3

The 1970 population was used as standard for age-adjustment.

Thus, "we are unable to say how much of the risks attributed to cigarettes is a 'pure' cigarette risk and how much is cigarette times another, possibly on the job hazard"[8]. Moreover, smoking and occupation are confounded variables, smoking among men being more prevalent in 'blue-collar' workers than in professional and managerial classes[13]. Occupational causes of lung cancer include asbestos, radon daughters, nickel ores, chromium, arsenic, beryllium, mustard gas, vinyl chloride and bischloromethyl ether, apart from incompletely identified carcinogens in a wide range of industries such as rubber curing, tanning, steel (coke ovens), foundries, automobile, and petrochemicals. Thus, lung cancer rates in asbestos insulation and topside coke oven workers are as much as 10 times greater than general population rates.

Underestimation of the role of such occupational carcinogens has been assisted by the fact that lung cancer mortality rates, based on the International Classification of Diseases, fail to distinguish pleural mesotheliomas from lung cancers; there is evidence of substantial under-reporting of mesotheliomas (by about 75%) in high risk groups[14], and even more so in occupations, such as automobile mechanics, where asbestos exposure has not been well recognized. There is a further lack of distinction between lung cancers of different histological types, some of which, such as adenocarcinomas, are less likely to be due to smoking than to occupational carcinogens[15,16]. "In several instances where the risk of bronchogenic carcinoma has been shown to be increased among occupationally exposed groups, there has been an accompanying shift in the distribution of histologic types of tumours", away from the small-cell undifferentiated and squamous cell carcinoma of the lung, the principal types whose frequency is increased by smoking, in the direction of other types, particularly adenocarcinoma[16]. "This (shift) has been noted among metal miners, uranium miners, copper smelter workers, vinyl chloride polymerization workers, chloromethyl methyl ether production workers, and mustard gas manufacturers" (ref. 16).

Possible variations of smoking patterns fail to account for the marked excess in US lung cancer rates identified in specific occupational exposures, particularly among ethnic minorities and migrants from southern states[17]. A further challenge to the dominant role ascribed to smoking seems to be provided by observations that the risk of lung cancer in certain occupational groups, such as American Indian uranium miners[18], Swedish zinc–lead miners[19], mustard gas workers[20], copper smelters exposed to arsenic[21], and chloromethyl methyl ether workers[22], is about as high among nonsmokers as smokers, although the latency period is reduced in smokers, suggesting a possible promotional effect of smoking. It appears that the relative risks of lung cancer for smokers as against risks for nonsmokers may have been overestimated, particularly in less than lifetime studies[13]. Variations in smoking do not account for geographic excesses in lung cancer rates in US males and females, which overall reflect proximity of residence to petrochemical and certain other industries[23,24]; there are also data showing associations between levels of atmospheric carcinogens and lung cancer mortality rates[25]. It may be noted that a report[26] from Peto's own institution demonstrates that the correlation coefficient between lung cancer and smoking internationally explains only one-third as much of the variation as does the correlation between lung cancer and solid fuel consumption (0.4 versus 0.7; $r^2 = 0.16$ versus 0.49).

Overemphasis on the carcinogenic effects of smoking, and ignoring or discounting the role of occupational and other exposures, is extended by Peto and others to cancers of the

Table 4 Changes in US cancer incidence rates from 1970 to 1975[7,8]

	Average % increase in incidence rates, 1970–75			
	Cancers of all sites		Cancers of all sites except lung	
Groups	Annual	5-Year	Annual	5-Year
White male	0.9	4.7	0.9	4.6
Non-white male	2.3	11.9	2.7	14.3
White female	2.2	11.6	1.8	10.2
Non-white female	6.1	34.6	5.7	32.2

Table 5 International correlations between breast, colon and liver cancers and possible aetiological variables[26]

	Correlation coefficients			
	Consumption of fat	Consumption of animal protein	Gross National Product	Total energy production
Breast cancer				
Incidence	0.79	0.77	0.83	0.70
Mortality	0.89	0.83	0.72	0.60
Colon cancer (M)				
Incidence	0.74	0.74	0.81	0.68
Mortality	0.85	0.86	0.77	0.69
Colon cancer (F)				
Incidence	0.78	0.80	0.82	0.67
Mortality	0.81	0.84	0.69	0.62
Liver cancer (M)				
Incidence	−0.49	−0.59	−0.42	−0.25*
Liver cancer (F)				
Incidence	−0.59	−0.67	−0.53	−0.31*

* 'Liquid energy'.

bladder and pancreas which are variously characterized as related to or caused by smoking[27,28]. However, the relative risks for these cancers are several times less in smokers compared with nonsmokers than is the case for lung cancer. Excess bladder cancer rates have been identified in several occupational categories, including rubber, paint manufacturing and textile dyeing workers[29], and among residents in highly industrialized counties[30], particularly those with large chemical industry complexes[31]. Excess pancreatic cancer rates have also been reported in various occupations including steel and metal workers[32] and organic chemists.

Recognition of the important role of occupational exposures in lung cancers previously ascribed, exclusively or largely, to smoking in no way detracts from the recognition, emphasized in *The Politics of Cancer*, that the impact of smoking constitutes a "national disaster". There is no basis for regarding the smoking/lifestyle and occupational theories as mutually exclusive, particularly as these exposures may operate interactively. Furthermore, lifestyle is a somewhat misleading rubric for smoking as it restrictively implies voluntary personal choice. Placing responsibility for personal choice of an addictive lethal habit on young teenagers, the fastest growing group of new smokers, seems inappropriate. Failure to control smoking reflects a wide range of political and economic constraints, including massive press advertising by the industry which omits the word 'death' from the guarded small print warning of danger, massive revenues to federal, state and local government from tobacco taxes, federal subsidies to the industry and unwillingness of governments to increase tobacco taxation or to develop incentives to tobacco farmers to diversify. It is also important that the industry has moved to open up massive new markets with high-tar cigarettes in less developed countries, where the population is poorly informed on the hazards of smoking.

Role of diet

Lifestyle proponents are on less sure ground when they bracket diet, excess fat and overnutrition with smoking as the causes of the majority of cancer deaths. This claim is based largely on international correlations between consumption of total fat and rates for cancer of the breast and colon[26]; however, such correlations by themselves are not proof of causality. Similar correlations were found, in the same study from Peto's institution, between breast and colon cancers and other variables, such as Gross National Product and consumption of animal protein, which also appear to reflect industrialization[26] (Table 5). Furthermore, "epidemiologically, the case against fat is weak because there are populations that have a high fat intake and little bowel cancer . . ."[33]. Of two case control studies on the association between diet and breast cancer, one found no effect[34] and the other found trivial effects of fat and caloric intake, concluding that ". . . recommendations of major dietary modification as a possible preventive measure for breast cancer are clearly premature"[35].

Equally unconvincing are the studies, cited by Peto as corroborative evidence on the experimental effects of diet, which were largely concerned with the influence of fat on the incidence of tumours induced by chemical carcinogens and ionizing radiation, and the influence of caloric intake on the incidence of spontaneous and induced tumours. Not only were different variables defined in the animal and human studies—per cent fat in the diet and total dietary fat, respectively—but increasing fat levels in the animal experiments were associated with increased incidence of skin, liver and breast cancers, whereas the reported correlations between fat consumption and liver cancer mortality are negative for both men and women (Table 5). Moreover, these experiments often failed to differentiate between variations of total dietary fat and caloric intake in test animals and to adjust caloric intake in controls to reflect dietary fat variations in test animals; the magnitude of the variations in fat and caloric intake required substantially to influence the incidence of induced and spontaneous tumours in experimental animals is generally far in excess of the dietary differences observed among the various human populations studied[33]. These experiments invariably failed to adjust the intake in controls of fat soluble carcinogens, present in fat as accidental environmental contaminants, to reflect variations of fat intake of test animals.

Peto's claim for the causal role of dietary fat in human cancer overstates the conclusions of those cited as the basis for his claims. Armstrong and Doll[26], for instance, merely suggest that dietary fat levels may influence the incidence of colon and breast cancers, without asserting causality. Doll considers that diet may act by modifying the incidence of tumours induced by carcinogens or by acting as a vehicle for exogenous carcinogens[36]—a suggestion also made in *The Politics of Cancer* which Peto dismisses as "implausible". Carrol concludes that "although caloric intake may be a factor in human carcinogenesis, it does not appear to offer a practical approach to the problem"[37]. As recognized by current concepts on the multifactorial aetiology of cancer, there is a substantial probability that a wide range of influences, diet and other lifestyle factors included, modify individual responses to carcinogenic agents. To ascribe causality to any particular modifying factor requires a degree of scientific evidence that has not yet been presented for dietary fat.

Role of occupation

Peto associates himself with the insistence by the chemical industry[2] and other lifestyle proponents that occupational exposures account for about 5% (refs 38–40) or "a very small proportion"[41] of all cancers. This view is based on ascribing given percentages to known or alleged lifestyle factors, including smoking, fatty diet and sunlight, leaving a small unaccounted for residue to which occupational factors are arbitrarily assigned by exclusion. The authors of this simplistic hypothesis compensate for its tenuous basis by reliance on 'educated estimates' and by making circular references to each other, often by 'personal communication', as the responsible authority.

However, there are problems with such 'guesstimates'. First, they fail to consider the multifactorial aetiology of cancer and the role of multiple causal agents, such as asbestos and smoking[42]; thus, the summation of known causes of cancer should properly exceed 100%. As one of the lifestyle authors recently stressed[36], "there is now strong evidence to suggest that the risk of cancer is commonly increased by interaction of two or more factors". Second, current cancer rates reflect exposures 20 to 30 years ago, when production levels of occupational carcinogens were a small fraction of the present; such estimates should thus now be adjusted to reflect increasing numbers of workers exposed. Third, the authors of these guesstimates failed to consider the very limited nature of the data base on exposure to occupational carcinogens. Nor have they at any stage protested or even commented on the persistent refusal of the chemical industry to make such critical data available. In the absence of exposure data, it is even less clear how the 'lifestylers' confidently arrive at their estimate of less than 5%.

Rather than addressing himself to such problems, Peto dismisses recent estimates of the importance of occupational carcinogens in a report by the US Public Health Service[43] as exaggerated, unsound and unreasonable. This report, prepared by nine named and internationally recognized experts in cancer epidemiology, statistics and carcinogenesis from three federal research agencies, is based on a National Occupational Hazard Survey which between 1972 and 1974 surveyed nearly 5,000 workplaces chosen to provide a cross-section of industry in the United States. The report estimated the total number of workers exposed to asbestos, nickel ores, chromium, arsenic, benzene and petroleum fractions, including aromatics. The excess cancers attributable to each of these carcinogens were derived by multiplying the number of exposed workers by known risk ratios and subtracting the "normal incidence" of the cancer.

The report concluded that "as much as 20% or more" of cancers in the near term and future may reflect past exposure to the six carcinogens considered. The uncertainties and limitations in these conclusions, including the possibility that exposures and risk ratios may have been overestimated in some instances, were clearly stated in the report, as were other considerations including the multifactorial aetiology of cancer, and the role of lifestyle factors and their possible interactions with occupational exposures.

The possibility that this government report underestimates rather than overestimates the role of occupational exposures, for several reasons some of which are recognized in the report, has not been considered by its denigrators, including Peto. First, the calculations in the report ignore the role of radiation and of some ten epidemiologically recognized occupational carcinogens, other than the six considered. Second, the risk ratios considered may be artificially low as they were largely derived from less-than-lifetime epidemiological studies, which may thus underestimate the true risk in view of the long latencies commonly involved. Third, the report does not consider the many statistical and methodological constraints common to most occupational epidemiological studies[44] such as relatively small numbers of workers in many locations, changes in exposure patterns over time due to employee turnover, plant shutdown, process and production changes and changes in management, all of which lead to fragmentation of health and exposure records, access to which is often restricted by industry. Fourth, the estimates fail to take account of the many chemicals recognized as carcinogenic in animals for which there are no exposure or epidemiological data. Thus, of 442 chemicals and industrial processes recently evaluated by the International Agency for Research on Cancer (IARC), epidemiological data are available for only 60 (14%), although evidence of experimental carcinogenicity was considered to be sufficient for 143 (32%)[45]. Fifth, the estimates exclude high risk occupations with incompletely defined carcinogens, such as the steel, rubber and tanning industries. Sixth, the estimates do not adequately reflect conditions in small business where exposure levels are likely to be higher than in major chemical companies. Seventh, the report does not reflect major increases in the production of the occupational carcinogens it considered such as benzene, with the likelihood of recently increasing exposures. Eighth, the study examined only a limited number of sites, excluding cancers such as skin and bladder which are known to be occupationally related. Finally, the estimates neglect the possible role of fugitive point-source emissions of industrial carcinogens as causes for the excess of overall and organ-specific cancers, including lung, bladder, colon, pancreas and breast, in residents of certain highly industrialized counties.

This government report has received extensive support from various expert bodies, such as the Toxic Substances Strategy Committee, whose position has been endorsed by 17 federal agencies, and international groups, such as the International Labor Organization, and the US and British trades union. The report has also received additional support in the critique of two consultants to the chemical industry's American Industrial Health Council which concluded that "... the full range (of total cancer attributable to occupational exposure) using multiple classifications may be from 10 to 33% or perhaps higher if we had better information on some other potentially carcinogenic substances.... The annual number of cancer deaths attributable to asbestos is in the range from 29,700 to 54,000, which corresponds to a percentage range of the total cancer of 7 to 14%.... Any argument over these numbers cannot detract from the fact that asbestos exposure was, as the authors (of the Government report) state, a major public health disaster.... We also believe that reduction of exposure to carcinogens in the course of employment can certainly be expected to affect major reductions in the frequencies of occurrence of cancer and is one of the most promising applications of preventive medicine"[47]. The American Industrial Health Council failed to release this critique until the record of the recent Occupational Safety and Health Administration hearings on regulation of occupational carcinogens closed.

Finally, there is no basis whatsoever for recent unsubstantiated allegations by Peto and others that all or most authors of the government report have disowned or rejected it or its conclusions (K. Bridbord, M. Schneiderman and A. Upton, personal communication). It should be further emphasized that this 50-page report was prepared as a government document specifically for inclusion in public hearing records, and not for submission to a scientific journal.

Conclusions

Cancer is a disease of multifactorial aetiology to which occupational exposure and smoking can contribute importantly, sometimes interactively. There have been substantial recent increases in cancer rates which cannot be accounted for by smoking alone. Smoking is the major lifestyle factor of importance in cancer, and evidence for the causal role of other lifestyle factors, particularly diet, is slender. The role of lifestyle factors has been exaggerated, by those with an economic or intellectual investment in this theory, by largely excluding involuntary exposures to carcinogens and minimizing the role of occupational carcinogens. These considerations further illustrate the primary thesis of *The Politics of Cancer*: cancer is essentially a preventable disease which requires intervention and regulation at several levels, particularly the occupational and smoking. Failure to prevent cancer reflects major political and economic constraints which have hitherto been largely unrecognized or discounted.

1. Peto, R. *Nature* **284,** 297–300 (1980).
2. Epstein, S. S. *The Politics of Cancer* (Sierra Club Books, San Francisco, 1978, revised and expanded in Anchor/Doubleday, New York, 1979); quotations refer to the 1979 edition.
3. *AIHC Recommended Alternatives to OSHA's Generic Carcinogen Proposal, Occupational Safety & Health Administration (OSHA) Docket No. H-090)* (American Industrial Health Council, 24 February, 1978).
4. *Vital Statistics—Special Reports, U.S. DHEW,* **43,** 163 (1956); *Vital Statistics of the U.S.* **2** (1970).
5. Pollack, E. S. & Horm, J. W. *J. natn. Cancer Inst.* **64,** 1091 (1980).
6. *Trends in Mortality, 1951–1975* (U.K. Office of Population Censuses and Surveys, Series DH1, No. 3, 1978).
7. *Third National Cancer Survey, 1969–1971 (NCI, Bethesda); Cancer Surveillance Epidemiology and End Results (SEER) Program.*
8. Schneiderman, M. *Occupational Safety & Health Administration OSHA Docket 090* (4 April, 1978).
9. Zdeb, M. S. *Am. J. Epidemiol.* **106,** 6 (1977).
10. Davis, D. L. & Magee, B. H. *Science* **206,** 1356 (1979); see also U.S. International Trade Commission Reports.
11. *Chemical Dangers in the Workplace*, 34th Report of the Committee on Government Operations, House of Representatives (September 27, 1976).
12. Enstrom, J. E. *J. natn. Cancer Inst.* **62,** 755 (1979).
13. Sterling, T. D. *Int. J. Hlth Services* **8,** 437 (1978).
14. Newhouse, M. L. & Wagner, J. C. *Br. J. ind. Med.* **26,** 302 (1969).
15. Wynder, E. L. & Stellman, S. D. *Cancer Res.* **37,** 4608 (1977).
16. Wagoner, J. K., Infante, P. F. & Bayliss, D. L. *Envir. Res.* **21,** 15 (1980).
17. Mancuso, T. F. & Sterling, T. D. *J. natn. med. Ass.* **67,** 107 (1975).
18. Archer, V. E., Gillam, J. D. & Wagoner, J. K. *Ann. N.Y. Acad. Sci.* **271,** 280 (1976).
19. Axelson, O. & Sundell, L. *Scand. J. Work envir. Hlth* **4,** 46 (1978).
20. Yamada, A. *Acta path. jap.* **13,** 131 (1963).
21. Pinto, S. S. *Archs envir. Hlth* **33,** 325 (1978).
22. Weiss, W. & Boucot, K. R. *J. Am. med. Ass.* **234,** 1139 (1975).
23. Mason, J. T. & McKay, F. W. *Atlas of Cancer Mortality for U.S. Counties: 1950–1969* (Dept H.E.W., Washington, D.C., 1975); *Atlas of Cancer Mortality for U.S. Counties among U.S. Non-Whites* (Dept H.E.W., Washington, D.C., 1976).
24. Blot, W. J. *et al. Science* **198,** 51 (1977).
25. Carnow, B. *Envir. Hlth Perspect.* **22,** 17 (1978).
26. Armstrong, B. K. & Doll, R. *Int. J. Cancer* **15,** 617 (1975).
27. Cole, P., Monson, R. R., Haning, H. & Friedell, G. H. *New Engl. J. Med.* **284,** 129 (1971).
28. Wynder, E. L., Mabuchi, K., Maruch, N. & Fortner, J. G. *Cancer* **31,** 641 (1973).
29. Cole, P. & Goldman, R. *Persons at High Risk of Cancer: An Approach to Cancer Etiology and Control* (ed. Fraumeni, J. F. Jr) 167–184 (Academic, New York, 1975).
30. Blot, W. J. & Fraumeni, J. *J. natn. Cancer Inst.* **61,** 1017 (1978).
31. Hoover, R. & Fraumeni, J. *Envir. Res.* **9,** 196 (1975).
32. Milham, J. in *Occupational Carcinogenesis* (eds Wagoner, J. & Saffiotti, U.) 243–249 (New York Academy of Sciences, 1976).
33. Berg, J. W. in *Persons at High Risk of Cancer: An Approach to Cancer Etiology and Control* (ed. Fraumeni, J. F. Jr) 201–224 (Academic, New York, 1975).
34. Phillips, R. L. *Cancer Res.* **35,** 3513 (1975).
35. Miller, A. B. *et al. Am. J. Epidemiol.* **107,** 499 (1978).
36. Doll, R. *Cancer Res.* **45,** 2475 (1980).
37. Caroll, K. K. *Cancer Res.* **35,** 3374 (1975).
38. Higginson, J. & Muir, C. S. *J. natn. Cancer Inst.* **63,** 1291 (1979).
39. Wynder, E. L. & Gori, G. B. *Am. J. Publ. Hlth* **66,** 359 (1977).
40. Maugh, T. H. *Science* **205,** 1363 (1979).
41. Doll, R. *Nature* **265,** 589 (1977).
42. Selikoff, I. J. & Hammond, C. E. *J. Am. med. Ass.* **242,** 458 (1979).
43. Bridbord, K. *et al. Estimates of the Fraction of Cancer in the United States Related to Occupational Factors* (National Cancer Institute, National Institute of Environmental Health Sciences, and National Institute for Occupational Safety and Health, 15 September, 1978).
44. Stellman, J. M. & Stellman, S. D. *What Proportion of Cancer is Attributable to Occupation: Statistical and Social Consideration* (American Society for Preventive Oncology, Chicago, 6–7 March, 1980).
45. Report of an IARC Working Group *Cancer Res.* **40,** 1 (1980).
46. *Toxic Chemicals and Public Protection* (Report to the President of the Toxic Substances Strategy Committee of the Council on Environmental Quality, May, 1980).
47. Stallones, R. A. & Downs, T. *A Critical Review of Bridbord, K. et al.* prepared for the American Industrial Health Council (1979).

ARTICLES

Proterozoic transcurrent tectonics: palaeomagnetic evidence from Venezuela and Africa

T. C. Onstott
Department of Geological and Geophysical Sciences, Princeton University, Princeton, New Jersey 08544
R. B. Hargraves
Direccion de Geologia, Ministerio de Energia y Minas, Caracas, Venezuela

Comparison of palaeomagnetic poles from the Guayana Shield with the published Precambrian palaeomagnetic record for Africa suggests that large-scale (thousands of kilometres) transcurrent motions have occurred along Upper Proterozoic mobile belts in South America and Africa. Displacements of such magnitude are consistent with a transform model for Precambrian mobile belts.

DISCUSSION of the role of plate tectonics in Precambrian crustal evolution has focused on the significance of the mobile belts that partition Precambrian shields (see Fig. 1). They could have been formed by intracratonic vertical movements and crustal thinning induced by mantle upwelling along hotspot tracks or ridge systems[1,2] (ridge-lineament model), or they may be the scars of the subduction zones along which the two adjacent cratons were 'welded' together[3] (suture model). The ridge-lineament (suture) model implies small (large) relative motions among cratonic nuclei.

It is also possible that Precambrian mobile belts are zones of transcurrent movement[4,5]. This hypothesis would predict differential motions of a different type among adjacent cratons. The suture model requires the strike of the mobile belt to fall approximately along a meridian (longitude) passing through the relative rotation pole of the adjacent cratons; whereas in the transcurrent model, the strike of the mobile belt will tend to lie along a latitude line with respect to the relative rotation pole.

With a sufficient amount of reliably dated palaeomagnetic data of the same age from each of the bordering cratons, it is theoretically possible to determine which of the three models best fits the data. We compare here newly obtained Precambrian palaeomagnetic data from the northern part of the Guayana Shield with the published Precambrian palaeomagnetic record for the West African, the Kalahari and the Congo shields. Although the palaeomagnetic data base is still rather small, we believe the data are more consistent with the transcurrent model than the other two models. Our results, however, as in other published intercratonic comparisons of Precambrian palaeomagnetic data[6–11], are not completely unambiguous.

Our palaeomagnetic data (Table 1) are based on a study of 300 samples (47 sites) from five major stratigraphic units of the Venezuelan Guayana Shield: the granulites of the Imataca Complex (2,000–2,050 Myr)[12,13], the Supamo–Pastora granite-greenstone terrain (1,900–2,000 Myr)[14–16], the Cuchivero-Caicara acid metavolcanics (1,600–1,900 Myr)[17–20], the Roraima sandstone and diabases (1.66)[21–23] and the Parguaza rapakivi batholith (1,520–1,590 Myr)[17].

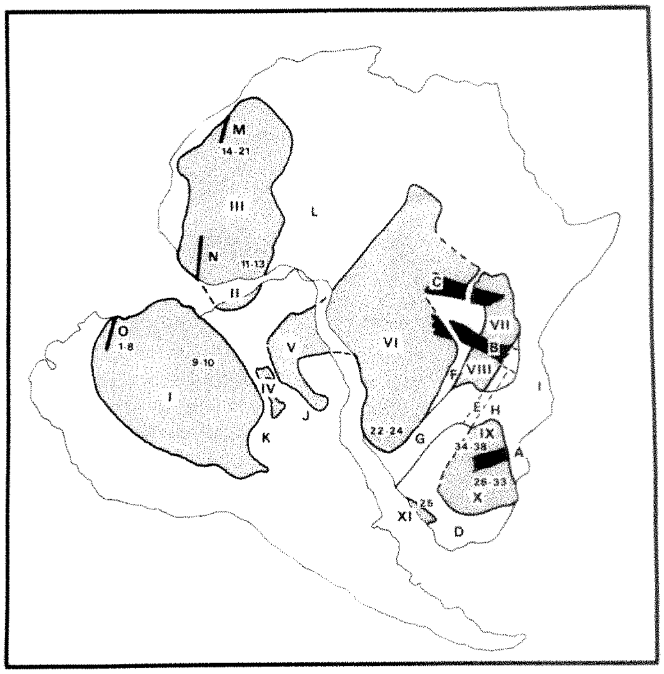


Fig. 1 Stable continental crust since 1,500 Myr BP, and mobile belts in Africa and South America. Plot is on the Bullard *et al.*[36] reconstruction. Stippled areas represent continental crust stable since 1,500 Myr. I, the Guayana Shield. The West African Shield is composed of the West African Craton (III) and the Sao Luis Craton (II). The Congo Shield is composed of the Sao Francisco Craton (V), the Kasai Craton (VI), the Tanzanian Craton (VII) and the Zambian Craton (VIII). The Kalahari Shield is composed of the Rhodesian Craton (IX) and the Kaapvaal Craton (X). XI is the Richtersveld Province, and IV is the Goias Central Massif. The Lower Proterozoic age (2,000 Myr.) mobile belts comprise the Limpopo belt (A), the Ubendian-Rusizian belt (B) and the Kibali-Toro-Bugandian belt (C). The Kibaran age (1,100 Myr) belts consist of the Namaqualand belt (D), the Irumide belt (E) and the Kibaride belt (F). The Pan African age (600 Myr) belts comprise the Damara belt (G), the Zambezi belt (H), the Mozambique belt (I), the Brasilia (Brazilian) belt (J), the Paraguay-Araguaia belt (K) and the Dahomeyan-Pharusian belt (L). M is the Sfariat zone, N is the Sassandra mylonite zone and O is the Guri mylonite zone. Arabic numerals refer to the pole numbers in the first column of Table 1.

Results of the analysis

Two or more components of magnetization were isolated by alternating field demagnetization from most of the sites. Thermal demagnetization of pilot samples of the Imataca Complex and the Roraima Group proved less effective than alternating field demagnetization in separating magnetic components. Examples of alternating field and thermal demagnetization results are given in Fig. 2. In general, the mean site directions of the more stable component as given by minimum dispersion analysis[24] were not significantly different from those given by vector analysis of Zijderveld diagrams[25,26]. The poles reported in Table 1 are averages of site poles based on at least four samples and with values for Fisher's $k > 7$. The mean site poles for the Imataca Complex and the Roraima Intrusive Suite were

0028-0836/81/030131—06$01.00

Table 1 Average palaeomagnetic poles for South America and Africa 1,500–2,200 Myr

Location–formation	Age (Myr)	No. of sites	Pole lat (N°)	Pole long (E°)	*k*	A95	Ref.
			Guayana Shield				
Venezuela							
1. Supamo–Pastora*	2,300–2,250‡ 1,900–2,000§	2	−50	346†	9	98	This work, 14–16
Imataca							
2. Pole I	2,055–2,000∥	9	−49	18†	34	9	This work, 12, 13
3. Pole II	2,055–2,000∥	4	−29	21	48	13	This work, 12,13
4. Cuchivero–Caicara	1,960–1,640∥	6	−69	70	10	22	This work, 17–20
5. Roraima sandstone	1,660∥	3	−69	347†	27	24	This work, 21
Roraima Intrusive Suite							
6. Pole I	1,660∥	9	−61	41	50	7	This work, 22, 23
7. Pole II		9	−44	353	54	7	60–62
8. Parguaza	1,590–1,524‡§∥	2	−67	26	407	13	This work, 17
Brazil							
9. Telles Pires	1,610∥	18¶	−69	292	—	9	39
10. Beneficente Gp.	1,480∥	21¶	−75	31	—	5	39
			West African Shield				
Ghana							
11. Obuasi greenstone	2,150–1,670§∥	5	−49	99	23	16	32, 33, 40
12. Obuasi dykes	2,150–1,670§∥	14¶	−56	69	22	—	32, 33, 40
13. Tarkwa dykes	2,150–1,670§∥	5	−53	36	32	14	32, 33, 40
Algeria							
14. Aftout gabbro	1,940∥	4	29	53	185	7	28, 31
15. Aftout gabbro	1,940∥	2	−6	97	8	>90	28, 31
16. Aftout diorite	1,940∥	3	38	45	45	19	28, 31
17. Aftout diorite	1,940∥	3	−20	80	65	15	28, 31
18. Aftout diorite	1,940∥	4	58	73	89	13	28, 31
19. Akilet dykes	<2,160∥	5	28	81	32	14	28, 31
20. Akilet dykes	<2,160∥	3	−16	54	52	17	28, 31
21. Akilet dykes	<2,160∥	13	−16	95	17	10	28, 31
			Sahara–Congo Shield				
Angola							
22. Cunene anorthosite	1,870–2,160∥	2	−26	16	18	63	29, 47
23. Cunene anotthosite	1,870–2,160∥	3	28	55	17	31	29, 47
24. Cunene anorthosite	1,870–2,160∥	6	−3	92	29	13	29, 47
			Richtersveld Province				
Namibia							
25. Orange River Lavas	2,000∥	5	−5	93	5	40	40, 42
			Kalahari Shield				
South Africa							
26. Lower. Ventersdorp Lavas	2,150	24¶	−3	42	11	9	43
27. Bushveld Complex	2,050∥	5	23	36	40	12	41, 44
28. Vredefort Complex	1,900∥	8	26	39	21	12	63, 64
29. Losberg Intrusion	1,840∥	3¶	33	37	182	9	
30. Palabora Complex	2,060‡	6	41	48	61	7	65
31. Palabora Complex		4	3	359	119	8	
32. Waterberg red beds	<1,790‡	4	2	334	29	17	66, 67
33. Waterberg red beds	<1,790‡	4	43	45	47	11	66, 67
Zimbabwe							
34. Great Dyke extensions	2,000∥	2	6	47	1,413	3	68, 69
35. Great Dyke extensions	2,000∥	2	−2	333	248	16	68, 69
36. Great Dyke extensions	2,000∥	11	29	4	45	7	68, 69
37. Mashonaland dolerites	1,810∥	14	13	341	51	6	45, 46

* Details on our site locations and on our palaeomagnetic data will be presented elsewhere[27].

† The palaeomagnetic data for the Imataca Complex, the Supamo–Pastora province and the Roraima Group have not been structurally corrected. The Imataca Complex fails the fold test at the 95% level of confidence. Although the between-site *k* values improve after structural corrections for the Supamo–Pastora data (from 6 to 12) and for the Roraima Group data (from 12 to 15), these are not signficant improvements at the 95% confidence level. For the Roraima diabases, the poles given here are the average of the data obtained in this study and previously published results[60–62].

‡ U–Pb ages.

§ K–Ar ages.

∥ Rb–Sr ages, in many cases whole-rock isochrons, $\lambda^{87}\text{Rb} = 1.42 \times 10^{-11}\ \text{yr}^{-1}$.

¶ Number represents total number of samples averaged.

distributed in clusters (separated by at least 120°). The mean and associated fisherian statistics for each cluster are reported (Table 1).

Although the protolith age[12] for the rock units of the Imataca Complex is >3,000 Myr, their stable remanence was probably acquired during regional uplift and cooling following granulite metamorphism, which has been dated at 2,050–2,000 Myr by a Rb–Sr biotite age[13] and detailed Rb–Sr and U–Pb analyses of the granulites[12].

The stable remanence in the Supamo–Pastora rock units was also probably acquired during regional uplift and cooling, which has been dated at 2,000–1,900 Myr based on K–Ar hornblende and biotite ages[15].

The origin of the stable remanence in the Cuchivero–Caicara metavolcanics is uncertain; however, because of their low metamorphic grade (lower greenschist) and lack of multiple components of magnetization, the age of their remanence is probably close to their Rb–Sr whole-rock isochron ages[17–20], which range from 1,960–1,640 Myr. The consistency in the remanence directions from sites separated by >400 km suggest little structural deformation has occurred since the acquisition of their magnetization.

Among the data from the Imataca Complex, the within-site concentration as given by Fisher's *k* varied widely from 6 to 255; the between-site concentrations were fairly high (Table 1). The within-site *k* values for the Supamo–Pastora and Caicara–Cuchivero units ranged from 10 to 165.

The more stable components present in the younger intrusive units, the Roraima diabases and the Parguaza batholith, are probably primary thermoremanent magnetizations acquired at 1,660 and 1,500 Myr. The within-site *k* values for the Roraima diabases were quite high, ranging from 17 to 230; those for the Parguaza were variable, ranging from 3 to 62.

The age and nature of the remanence of the Roraima sandstone are uncertain because hydrothermal alteration is pervasive and the remanence could easily be a chemical overprint.

Many of the less stable secondary directions observed in both the metamorphic and intrusive units represent partial remagnetization during the low-grade Nickorie metamorphic episode[15] dated at 1,200 Myr. Eight of the sites were apparently partially remagnetized during this metamorphic episode. These secondary overprint directions have not been included in Table 1, but will be discussed elsewhere[27].

Palaeomagnetic data

The Precambrian palaeomagnetic poles from the three major African cratons given in Table 1 have been recalculated from the original site–pole distributions. In most cases, the pole obtained is the same as that reported in the original publication. In the data from studies in Algeria[28], Angola[29] and Namibia[30], however, the site–pole distribution showed several distinct clusters. We have calculated poles for each of these clusters rather than a single site-mean pole as originally reported. We believe that the several clusters may be records of the geomagnetic field at different times, and that information is lost by calculating only one mean pole position. Our conclusions, however, do not depend on this operation, and the original published poles could equally well have been used.

All of the African palaeomagnetic data reported in Table 1 and Figs 3, 4 and 5 are based on at least 10 independent samples and have a Fisher's $k \geq 5$ (only three poles have $k < 10$).

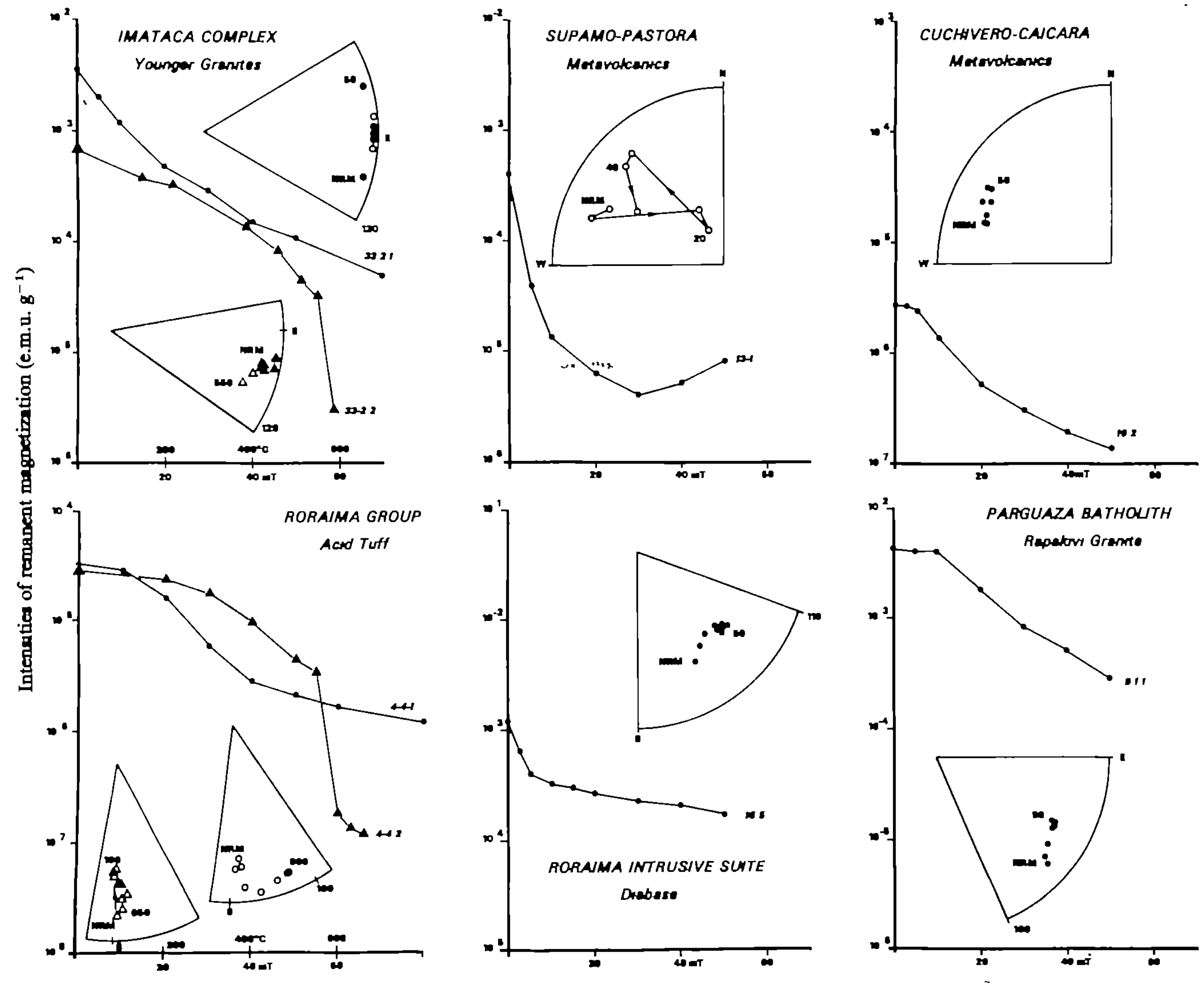


Fig. 2 Progressive demagnetization studies. Representative examples of thermal (▲) and alternating field (●) demagnetization of pairs of specimens from the same oriented samples of the Imataca Complex and the Roraima group, and representative examples of alternating field (●) demagnetization for specimens of the Pastora metavolcanics, Cuchivero metavolcanics, the Roraima diabase and the Parguaza batholith. The directions of remanent magnetization are plotted on equal-area projections, the plane is the present horizontal. Amphibolites and granites intrusive into the Imataca Complex tend to have very stable single-component magnetizations, whereas, mafic granulites, charnockites and biotite gneisses tend to have multicomponent magnetizations. Blocking temperatures indicate that magnetite is the principal carrier of the remanence, in polished sections magnetite is frequently present as very fine rods in pyroxene grains. The remanent magnetization of the greenschist facies rock units of the Pastora tends to be very weak and migrate widely on demagnetization. The remanent magnetization of the very low-grade Cuchivero metavolcanics and the tuffs of the Roraima group tends to be weak, but extremely stable. Blocking temperatures suggest that the magnetization is primarily carried by magnetite, which is present in the groundmass as very fine equant grains. The remanent directions tend to migrate rapidly to a stable, characteristic direction with alternating field demagnetization. The secondary directions seem to be lightning induced. Samples of the Parguaza rapakivi granite exhibited a wide range of demagnetization behaviour, from extremely high coercive single-component magnetization to low coercive multicomponent magnetizations.

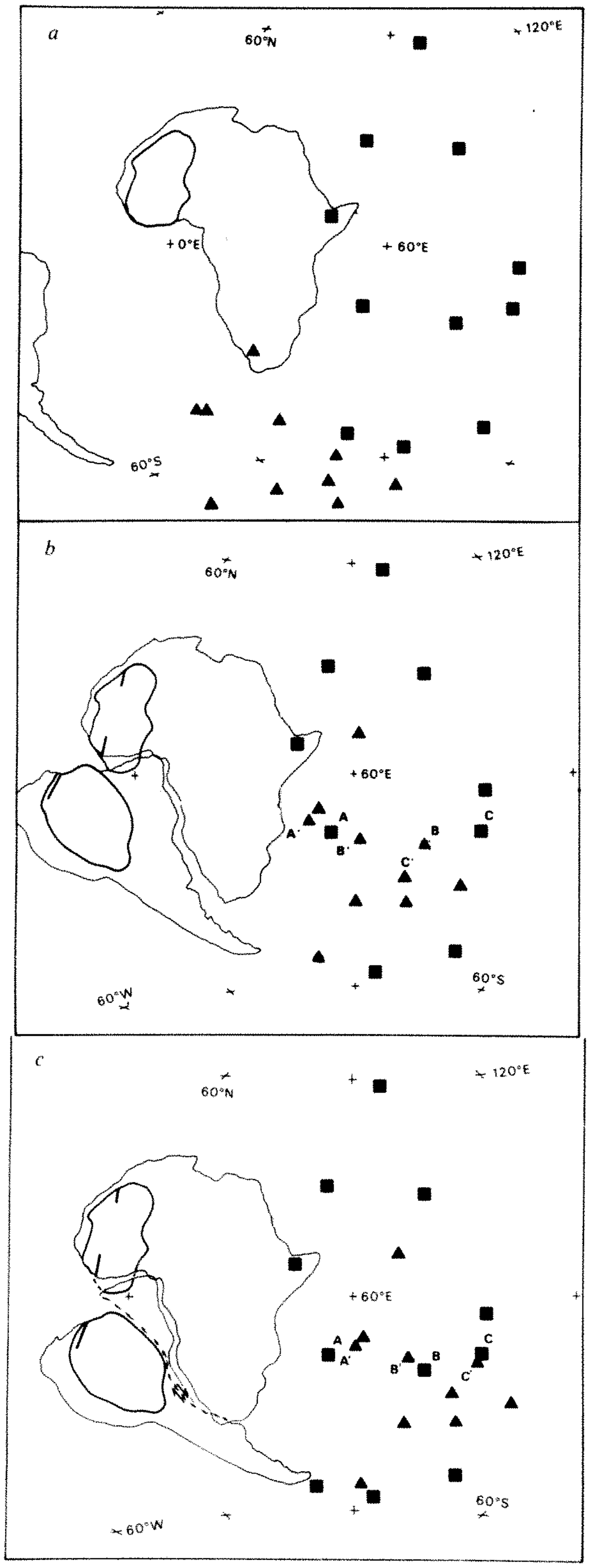


Fig. 3 Palaeomagnetic data from the South American Guayana Shield (▲) and the West African Shield (■) for 1,500–2,100 Myr. Continental crust older than 1,500 Myr and the Guri and Sassandra fault zones are shown (thick lines). *a*, Continents in their present day position. *b*, Reconstruction according to ref. 36, with A, B and C and A', B' and C' indicating the more reliably dated poles from the West African Shield and the Guayana Shield, respectively. *c*, Proposed reconstruction for the period 2,100–1,500 Myr, which brings together the correlatable data points from each shield and also aligns the basement structures between the two shields. Transcurrent fault is shown with direction of displacement indicated by small arrows.

Radiometric ages given in Table 1 are based on data given in refs 15, 16.

The West African data are from a sequence of rock units geologically similar to those in the Guayana shield: a metamorphic basement (1,800–2,000 Myr)[31–33] with intrusive diabases (1,300–1,700 Myr)[28,34]. Comparison of Fig. 2*a* and *b* shows that the South American data (triangles) do agree much better with the West African data (squares) after the South Atlantic has been closed using a Bullard *et al.*[36] fit, in agreement with the hypothesis of the Hurley and Rand[35].

Detailed comparison of one of the Imataca granulites and both of the Roraima diabase poles (numbers 3, 6 and 7 marked A', B' and C' in Fig. 3*b*, *c*) with the most likely correlative poles from the West African Shield, the Aftout diorite pole and the Akilet dyke poles (numbers 17, 20 and 21 marked A, B and C in Fig. 3*b*, *c*) suggests a slight westward offset in the South American data. This offset could be eliminated by rotating the Guayana Shield (Fig. 3*c*) 13° counterclockwise about a pole 70°E, 66°N. Such a rotation would require ~1,500 km displacement along a fault zone postulated to lie in the Brazilian mobile belt (Fig. 1). After this shift, the Sassandra mylonite zone[37] (Fig. 1), which separates the granulite rock units of Ivory Coast from the Birrimian greenstones, becomes continuous with the Guri mylonite zone which separates the Imataca granulites from the Pastora greenstones[38] (Fig. 3*c*). This alignment suggests that the proposed translations may be correct.

In comparing the data from the Guayana Shield and the West African Shield with those from the Congo Shield and the Kalahari Shield (Figs 4, 5), we have shown only those data that fall in the 1,800–2,100-Myr age range, because no palaeomagnetic data in the 1,500–1,800-Myr age range exist for the latter two shields. In Fig. 4 the three poles from Ghana (11–13 in Table 1, Fig. 1) are significantly removed from other contemporaneous West African poles as well as from the poles of the other shields. As observed in Fig. 3*b*, *c*, the Ghana poles do coincide with a pole from Brazil, for which ages of 1,500–1,600 Myr have been reported[39]. This correlation leads us to suspect that the age of the magnetization for the poles from Ghana may be younger than the 2,200-Myr age given for them[40]. For this reason and to clarify the diagram, all the Ghana poles have been omitted from Fig. 5.

The principal observation from Fig. 4 is that the palaeomagnetic data from the Kalahari Shield fall mainly in eastern and northern Africa. In contrast the palaeomagnetic poles from all the other cratons are concentrated in the Indian Ocean. The most recent age determinations[12,41,42] for both data sets strongly suggest that this discrepancy cannot be the result of differences in age.

The Kalahari palaeomagnetic poles exhibit a consistent succession in age from the Lower Ventersdorp Lavas pole (number 26, 2,150 Myr)[43], through the Bushveld Complex pole (number 27, 2,000 Myr)[41,44], to the Mashonaland dolerites pole (number 33, 1,800 Myr)[45,46]. The data from the Cunene Anorthosite Complex (numbers 22–24, 1,800–2,100 Myr[29,47]; diamonds in Figs 4 and 5) of the Congo craton and from the Orange River Lavas (number 25, 2,000 Myr[30,42]; solid hexagon) of the recently recognized Richtersveld Province in Namibia[48] (Fig. 1) are distinct from those of the Kalahari, and compare favourably with poles from the West African and Guayana shields; a temporal distribution pattern for these latter poles is roughly comparable with that displayed by the Kalahari poles (Fig. 4). The best superposition of these two apparent polar wander curves as drawn is obtained by rotating the Guayana, West African and Congo shields and the Richtersveld Province 64° counterclockwise about a pole at 44°E, 28°S (Fig. 5).

As the Richtersveld Province (XI in Fig. 1) also takes part in the rotation, a corresponding zone of right lateral transcurrent displacement must separate it from the Kalahari Shield (IX and X in Fig. 1). The only possible location for this zone is in the 1,100-Myr old Namaqualand mobile belt (D in Fig. 1). Similarly, the Kasai craton (VI in Fig. 1) must be separated from the Kalahari Shield by the northward continuation of this zone of dislocation, which would presumably lie within either the Irumide or the Kibaran mobile belt (E and F, respectively, in Fig. 1), as they are equivalent in age to the Namaqualand belt. If the rotation is attributed solely to transcurrent motion, then a small circle about the rotation pole (P in Fig. 5), passing through the Namaqualand belt, coincides more closely with the Irumide belt than with the Kibaran belt. The rotation would require ~2,000 km of right-lateral displacement within this combined Namaqualand–Irumide mobile belt. As no palaeomagnetic data exist for the Zambian–Tanzanian cratons for this time period, we cannot preclude the possibility that they remained fixed with respect to the Kalahari Shield, and that the displacement occurred along the Kibaran belt. Note, however, that after our postulated rotation, along the Irumide belt, the Limpopo belt (A in Fig. 1) and the Ubendian–Rusizian belt (B in Fig. 1) become continuous (Fig. 5). We consider this result to support primary displacement occurring along the Irumide belt.

Discussion

The geology of Southern Africa is generally consistent with our hypothesis except in the critical area south-east of the Zambian craton where it is extremely complex and much disputed[70]. Ramsay and Ridgway[49] find no evidence for an Irumide belt, and propose that the Pan African Mozambique belt forms the eastern margin of the Zambian (and Tanzanian) cratons. They do admit, however, that older gneisses of Ubendian or Kibaran (Irumide) age may be present within the Mozambique belt. Shackleton[50] recognizes an Irumide belt, but proposes that it terminates south-west of the Ubendian belt, which he concludes can be connected to the eastern end of the Limpopo belt through a trend of granulites in the Mozambique belt. In addition, Key and Sutton[51] propose that the Limpopo belt does not continue westwards completely through the Kalahari shield (Fig. 1), implying that its alignment with the Ubendian belt after rotation (Fig. 5) is fortuitous. Nevertheless, the palaeomagnetic evidence for our hypothesis is compelling, and hopefully any apparent inconsistency with the geology may eventually be resolved.

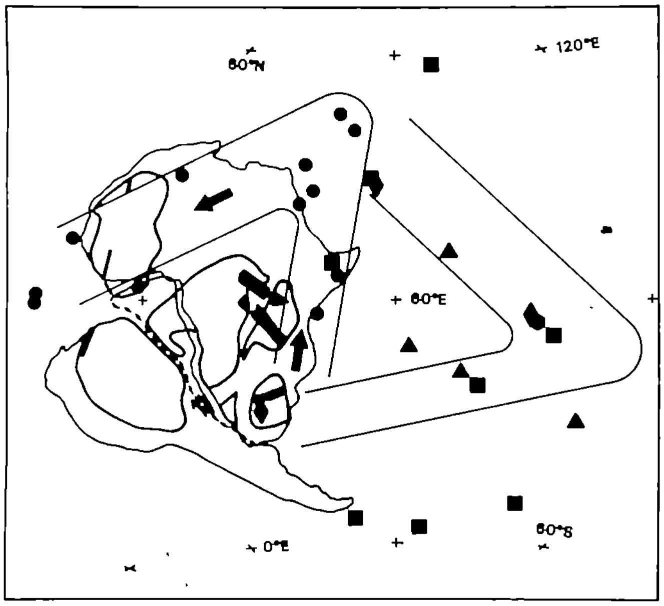


Fig. 4 Palaeomagnetic data for the Kalahari Shield (●), the Congo Shield (◆), the Orange River Lavas (⬢), the West African Shield (■) and the Venezuelan Guayana Shield (▲) for 1,800–2,100 Myr. Suggested apparent polar wander (APW) path for the Kalahari data is outlined and the direction of polar wander indicated (black arrows). Proposed correlative APW for the data from the other shields is outlined. The three solid squares near the bottom of the figure are the data from Ghana[40] and have not been included in the proposed APW path (see text). Continental crust older than 1,500 Myr (enclosed by thick lines), correlative fault zones (thick lines) and correlative mobile belts (solid) are shown. Continents are in the same reconstruction as Fig 3*c*

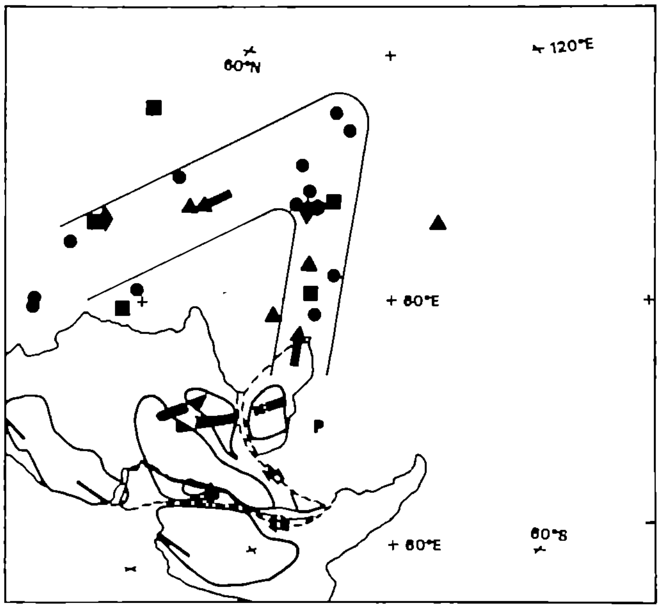


Fig. 5 Palaeomagnetic data same as in Fig. 4, except the data from Ghana have not been included. Continents are in proposed reconstruction, which brings together the two APW paths and aligns the Limpopo belt and the Ubendian–Rusizian belt. P is the pole of rotation, and the transform fault zone is shown with displacement indicated by small arrows. See Fig. 4 for remaining symbols.

Large-scale (thousands of kilometres) transcurrent displacements along Phanerozoic orogenic belts have been documented palaeomagnetically and geologically[52–55]. The above palaeomagnetic analyses suggest that large-scale transcurrent motions have also occurred in Proterozoic mobile belts, particularly those for which intense shearing and smaller-scale transcurrent displacements have been reported[56–59]. Without proper corrections for such movements polar wander paths based on an expanding intercratonic palaeomagnetic data base either will continue to become more complicated and convoluted, or the path widths and the apparent uncertainties in the ages of magnetization will be enlarged when attempts are made to keep the polar wander paths relatively simple.

If palaeomagnetic evidence is found for similar intercratonic movements in other continents then the transcurrent-tectonic model advocated here may signify a tectonic style in the Precambrian similar to that proposed by Sutton and Watson[10], which was characterized more by transcurrent shearing (jostling) of cratonic blocks than by the rifting, spreading and closing of oceans.

This research was supported by the Direccion de Geologia, Ministerio de Energia y Minas. We thank Drs W. J. Morgan, D. Jurdy, F. B. VanHouten and E. Irving for helpful criticisms and also N. Dorety and L. Hollingshead for technical assistance. The laboratory work was supported by the NSF grant EAR 76-22644 AO1.

Received 7 July, accepted 22 October 1980.

1. Hurley, P M in *Implications of Continental Drift to the Earth Sciences 2* (eds Tarling, D H & Runcorn, S K) 1083–1089 (Academic, London, 1973)
2. Kroner, A. *Precamb Res* 4, 163–213 (1977)
3 Burke, K. C & Dewey, J F in *African Geology* (eds Dessauvagie, T F J. & Whiteman, A J) 583–608 (University of Ibadan, Nigeria, 1970)
4 Sutton, J. & Watson, J. V *Nature* 274, 433–435 (1974)
5 Katz, M. B. in *The Early History of the Earth* (ed Windley, B F) 147–158 (Wiley, New York, 1976)
6 Irving, E & McGlynn, J. C *Geophys J R. astr. Soc* 58, 309–336 (1979)
7 Piper, J D A, Briden, J. C & Lomax, K *Nature* 245, 244–248 (1973)
8 Briden, J. C *Phil Trans R. Soc.* A280, 405–416 (1976)
9 McElhinny, M W & McWilliams, M. O *Tectonophysics* 40, 137–159 (1977)
10 Burke, K., Dewey, J F. & Kidd, W S F. *Tectonophysics* 33, 287–299 (1976)
11. Cavanaugh, M D & Seyfert, C K. *Geology* 5, 207–211 (1977).
12 Montgomery, C V & Hurley, P. M. *Earth planet Sci. Lett.* 39, 281–290 (1978).
13. Short, K C & Steenken, W F. *Asoc Venezolana Geol, Miner Petrol Boln Inf* 5, 189–221 (1962)
14 Olmeta, M. A *Boln Geol, Caracas* 10, 345–356 (1968).
15 Klar, G. thesis, Case Western Reserve Univ. (1979)
16. Gibbs, A. K. & Olszewski, W J. Jr *Zircon U–Pb Ages of Guayana Greenstone–Gneiss Terrain* (in the press)

17. Gaudette, H. D., Mendoza, V., Hurley, P. M. & Fairburn, H. W. *Bull. geol. Soc. Am.* **89,** 1335–1340 (1978).
18. Priem, H. N. A., Boelrijk, N. A. I. M., Hebeda, E. H., Verdurmen, E. A. Th. & Verschure, R. H. *Bull. geol. Soc. Am.* **82,** 1667–1680 (1971).
19. Amaral, G. & Halpern, M. in *X Inter-Guianas Geological Conf.* 474–487 (Belem, Brazil, 1976).
20. Basei, M. A. S. & Teixeira, W. *X Inter-Guianas Geological Conf.* 453–486 (Belem, Brazil, 1976).
21. Priem, H. N. A., Boelrijk, N. A. I. M., Hebeda, E. H., Verdurmen, E. A. Th. & Verschure, R. H. *Bull. geol. Soc. Am.* **84,** 1677–1684 (1973).
22. Hebeda, E. H., Boelrijk, N. A. I. M., Priem, H. N. A., Verdurmen, E. A. Th. & Verschure, R. H. *Earth planet. Sci. Lett.* **20,** 189–200 (1973).
23. Snelling, N. J. & McConnell, R. B. *Geologie Mijnb.* **48,** 201–213 (1969).
24. Irving, E., Stott, P. M. & Ward, M. A. *Phil. Mag.* **6,** 225–241 (1961).
25. Hargraves, R. B. & Burt, D. M. *Can. J. Earth Sci.* **4,** 357–369 (1967).
26. Dunlop, D. J. *Phys. Earth planet. Inter.* **20,** 12–24 (1979).
27. Onstott, T. C. & Hargraves, R. B. *9th Caribbean Conf. Volume (Dominican Republic)* (in the press).
28. Sabate, P. & Lomax, K. *Bull. B.R.G.M. (2)* **4,** 292–311 (1975).
29. Piper, J. D. A. *Phys. Earth planet. Inter.* **9,** 353–363 (1975).
30. Piper, J. D. A. *Geophys. J. R. astr. Soc.* **40,** 313–334 (1975).
31. Lassere, M., Lameyere, J. & Buffiere, J. M. *Bull. B.R.G.M. (2)* **4,** 5–13 (1970).
32. Bonhomme, M. *Annls Fac. Sci. Univ. Clermont* **5** (1962).
33. Tugarinov, A. L. *Geochimiya* **11,** 1336–1349 (1967).
34. Knopf, D. *Soc. Dev. Min. Cote d'Ivoire Mem.* 188 (1967).
35. Hurley, P. M. & Rand, J. R. in *The Ocean Basins and Margins* (eds Nairn, A. E. M. & Stehli, F. G.) **1,** 391–410 (Plenum, New York, 1973).
36. Bullard, E. C., Everitt, J. E. & Smith, A. G. *Phil. Trans. R. Soc.* A**258,** 41–75 (1965).
37. Dillon, W. P. & Sougy, J. M. A. in *Implications of Continental Drift to the Earth Sciences* Vol. 2 (eds Tarling, D. H. & Runcorn, S. K.) 1091–1095 (Academic, London, 1973).
38. Kalliokoski, J. *Bull. geol. Soc. Am.* **76,** 1027–1050 (1965).
39. Pacca, I. G. & Cordani, U. G. in *Prog. Abstr. XVII IUGG General Assembly* (ed. Fukushima, N.) 181 (Canberra, Australia, 1979).
40. Piper, J. D. A. & Lomax, K. *Geophys. J. R. astr. Soc.* **34,** 435–450 (1973).
41. Hamilton, *J. Petrol.* **18,** 24–52 (1977).
42. Reid, D. L. *Earth planet. Sci. Lett.* **42,** 368–378 (1979).
43. Henthorn, D. I. *17th A. Rep. Res. Inst. Afr. Geol.* 53 (University of Leeds, 1973).
44. Gough, D. I. & van Niekerk, C. B. *Phil. Mag.* **4,** 126–136 (1959).
45. Compston, W. & McElhinny, M. W. *Precamb. Res.* **2,** 305–315 (1975).
46. McElhinny, M. W. & Opdyke, N. D. *J. geophys. Res.* **69,** 2465–2475 (1964).
47. Da Silva, A. T. S. F. *Serv. Geol. Min. Angola Mem.* **13,** 1–10 (1974).
48. Blignault, H. J. in *Bull. Precambrian Res. Unit, Univ. Cape Town,* **15,** 49–56 (1974).
49. Ramsay, C. R. & Ridgway, J. *Precamb. Res.* **4,** 321–337 (1977).
50. Shackleton, R. M. *Implications of Continental Drift to the Earth Sciences* Vol. 2 (eds Tarling, D. H. & Runcorn, S. K.) 1083–1089 (Academic, London, 1973).
51. Key, R. M. & Sutton, R. M. *Precamb. Res.* **3,** 79–91 (1976).
52. Irving, E. *Can. J. Earth Sci.* **16,** 669–694 (1978).
53. Harland, W. B. *Geology* **8,** 258–259 (1980).
54. Van der Voo, R. *Geology* **8,** 259–260 (1980).
55. Morel, P. & Irving, E. *J. Geol.* **86,** 535–561 (1978).
56. Sutton, J., Watson, J. V. & James, T. C. *Tanganyika Bull. geol. Surv. Tangany.* 22 (1954).
57. Kroner, A. *Phil. Trans. R. Soc.* A**280,** 541–553 (1976).
58. Beach, A. *Phil. Trans. R. Soc.* A**280,** 569–604 (1976).
59. Dunnet, D. *Phil. Trans. R. Soc.* A**280,** 641–654 (1976).
60. Hargraves, R. B. *Geophys. J. R. astr. Soc.* **16,** 147–160 (1968).
61. Veldkamp, J., Mulder, F. G. & Zijderveld, J. D. A. *Phys. Earth planet. Inter.* **4,** 370–380 (1971).
62. Hargraves, R. B. *Phys. Earth. planet. Inter.* **16,** 277–284 (1978).
63. Hargraves, R. B. *J. Geol.* **78,** 253–263 (1970).
64. Davies, R. D., Allsopp, H. L., Erlank, A. J. & Manton, W. T. in *Spec. Publ. geol. Soc. S. Afr.* **1,** 576–593 (1970).
65. Briden, J. C. *Phil. Trans. R. Soc.* A**280,** 405–416 (1976).
66. Jones, D. L. & McElhinny, M. W. *J. Geophys. Res.* **72,** 4171–4179 (1967).
67. Oosthuyzen, E. J. thesis, Univ. Pretoria (1964).
68. Jones, D. L., Robertson, I. D. M. & McFadden, P. L. *Trans. geol. Soc. S. Afr.* **78,** 57–65 (1975).
69. van Breemen, O. & Dodson, M. H. *Bull. geol. Soc. Am.* **83,** 2005–2018 (1972).
70. Cahen, L. & Snelling, N. J. *The Geochronology of Equatorial Africa* (North-Holland, Amsterdam, 1966).

The *b*-value as an earthquake precursor

Warwick D. Smith

Geophysics Division, Department of Scientific and Industrial Research, Wellington, New Zealand

A systematic study of b*-values in New Zealand has shown that within the vicinity of forthcoming large earthquakes there is initially an increase in* b *and then a return to normal. The Caracas and San Fernando earthquakes show the same phenomenon, which is related to that of seismic gaps.*

IN the recurrence relation of Gutenberg and Richter[1], $\log N = a - bM$, the parameter b measures the relative numbers of small and large earthquakes. Values near unity are common. The value of b is usually calculated by the method of maximum likelihood (see ref. 2) whence

$$b = \frac{\log_{10} e}{\bar{M} - M_0}$$

$\bar{M}$ is the mean magnitude of the sample and M_0 the minimum magnitude considered. The standard error in b is $b/\sqrt{n}$ for a sample of n earthquakes[2] and the 95% confidence limits are twice this value.

Fiedler[3] found that *b*-values increased before the Caracas earthquake of 1967. His estimates were of the 'extreme *b*-value' for each year, and they increased to 1.33 before dropping again before the earthquake. Wyss and Lee[4] considered a sample of many more earthquakes, of much smaller magnitudes, in California. They calculated *b* using the maximum likelihood method and a 50-event window, and claim that it decreased both before and after main events. In their data the decrease before the main events took place over some hundreds of days. But there is also a suggestion of an increase in *b*, followed by a decrease just before the main event, the changes taking place over a few tens of days (for example at Danville, late 1971, where the magnitude of the main shock was 3.4).

Studies of aftershock sequences in New Zealand have also found that the *b*-value is not constant in time. In some cases Gibowicz[5] found an increase after the main event, followed by a decrease before the next large shock. He also found large changes in *b* for large areas of the North Island[6], as did Robinson[7], who found an increase to a peak value of 1.75 in 1971, followed by a decrease.

This paper examines further evidence relevant to Fiedler's suggestion[3] that a period of high *b*-values can occur before an earthquake.

High *b*-values in small regions

A systematic study of the temporal variation of *b*-values throughout New Zealand has now been made, using regions 2° in latitude by 2° in longitude, earthquakes of magnitude 4.0 and greater from 1955 to June 1979, focal depths less than 40 km, and a 50-event window which slides by 10 events. Smaller regions did not contain enough data for adequate resolution of the *b*-values. No attempt has been made to remove swarms and aftershocks from the data. Results for several of the regions are shown in Fig. 1, where the *b*-value is plotted at the time the last event entered the window. All five plots in Fig. 1 begin with low values of *b*, and this implies a lack of detection of small events before the mid 1960s when the New Zealand seismograph network was expanded to reduce the detection threshold to

Table 1 Total precursor times, from the onset of the increase in *b*-values to the occurrence of the earthquake

Earthquake	Magnitude	Precursor time
Tejon Pass, 1961	5.0	3.5 yr
Caracas, 1967	6.7	7 yr
Inangahua, 1968	6.8	6 yr
San Fernando, 1971	6.4	7 yr
Region *E*, 1971	5.9	3.2 yr
Region *C*, 1973	5.7	4 yr
Region *A*, 1977	6.0	7 yr
Pukaki, 1978	3.8	6 months
Pukaki, 1978	4.6	16 months

0028-0836/81/030136—04$01.00

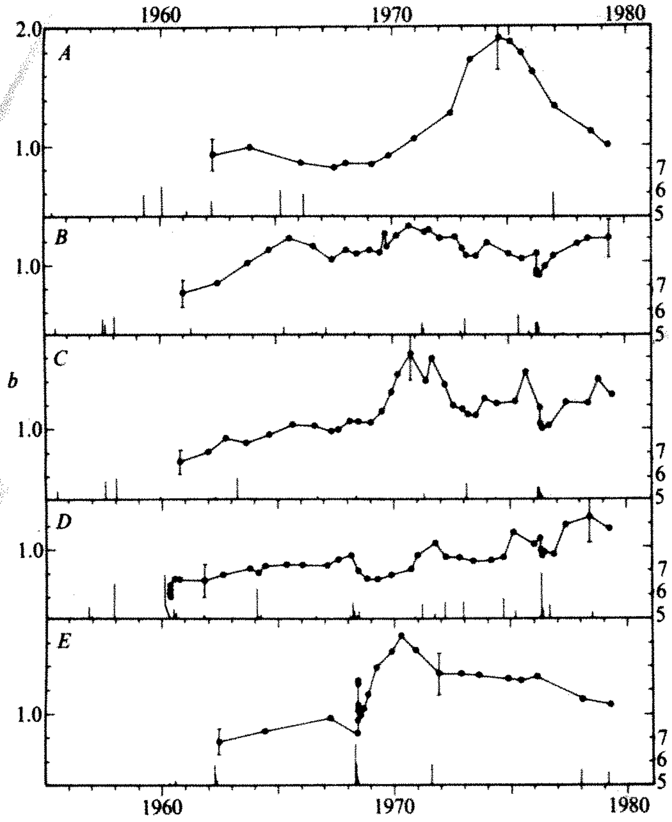


Fig. 1 The variation of parameter *b* from 1955 to 1979, in each of five regions of New Zealand. A location map is given in Fig. 2. Standard errors are proportional to *b* and are shown in a few cases, as are earthquakes of magnitude greater than 5. A 50-event window has been used, sliding by 10 events.

magnitude 4.0 throughout the country. Standard errors, proportional to the ordinate because of the uniform window length, are shown in several cases. Also shown are the earthquakes of magnitude >5 in each area. A location map is shown in Fig. 2.

In region *A* the *b*-value rose to a peak of 1.9 in 1974 and then decreased. During the decrease there was an earthquake of magnitude 6.0. The high *b*-values were accompanied by a paucity of moderate earthquakes. During the period 1969–76, while the *b*-value rose to its peak then began to drop, there were no earthquakes over magnitude 5.0 in the region.

Figure 1*B* shows the *b*-value variation for another area, in the east of the North Island. Little variation is evident, although there was an earthquake of magnitude 5.9 in June 1975. Figure 1*C* plots a region in the north-east, where a high value of *b* (1.6) was followed by an earthquake of magnitude 5.7 in 1973 which occurred near the city of Napier. Figure 1*D* shows a similar plot for the Fiordland area, exhibiting very little variation in *b* in spite of the large earthquake in 1976. This point is taken up again later.

Figure 1*E* shows the Inangahua earthquake of 1968 (magnitude 6.8), and a marked increase in *b* which followed it. It is clear that high values of *b*, while they may be associated with imminent earthquakes, may also relate to previous activity. If Fig. 1*E* is any indication, however, the previous event should be easily identified, and in this case the period of high *b*-values could also be associated with the magnitude 5.9 event in 1971.

Figure 3 shows data obtained from the Lake Pukaki microearthquake network, situated in the South Island of New Zealand near 44°S, 170°E. This network of nine stations was installed in 1975 to monitor seismicity there during the raising of the lake level for hydroelectricity generation. *b*-Values have been calculated for magnitudes 1.5 and greater, using a 50-event window which slides by 10 events. In December 1978 an earthquake of magnitude 4.6 occurred close to the centre of the network[8]. This event was preceded by a period of high *b*-values. There may also be another occurrence of the same phenomenon before the magnitude 3.8 shock of August 1978, superimposed on the larger trend.

Further evidence concerning whether or not periods of high *b*-value are followed by large earthquakes must be sought in other seismic regions, by systematic studies such as has been done in New Zealand. But the converse—whether or not large earthquakes are preceded by periods of high *b*-value—is examined in Fig. 4. The region 34–35°N, 118–119°W includes the San Fernando earthquake of 1971. Data selected from the Southern California catalogue[9] are shown in Fig. 4*A*. The magnitude threshold is 2.5 and the window length is 20 events, sliding by five events. Decreasing the window length was necessary because of the small area chosen and the generally low level of seismicity, even at the magnitude 2.5 level, but, of course, the standard error increases. Two periods of high *b*-value are evident (approximately double the background level), one before the earthquakes of 1961 and 1963 in the Tejon Pass quadrangle, and the other before the San Fernando earthquake of 1971. The data set did not include the San Fernando earthquake itself.

Figure 4*B* shows a reexamination of the data for area *e* (see Figs 1 and 2). The window length has been reduced to 20 events, sliding by five events. The *b*-value increased to 2.2 before the Inangahua earthquake of 1968. This is an alternative expression of the precursory gap which Evison[10] found for that earthquake. The area is of only moderate seismicity, and detection may not have been complete to magnitude 4.0 in the early 1960s. Both of these reasons demanded the short window. There is no clear decrease in *b* before the earthquake, but note that the final value of 2.2 at the beginning of 1968 represents an average back to 1963. There is clearly insufficient data to resolve any decrease which may have occurred before the major event. Extending the data sets to include the San Fernando and Inangahua earthquakes returns the *b*-value to its background level in each case.

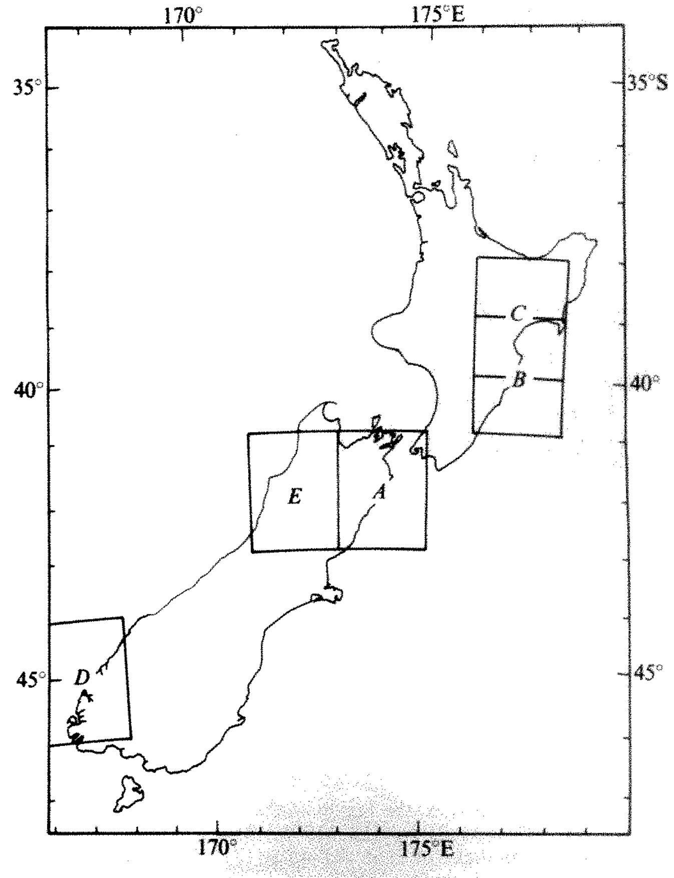


Fig. 2 Location map for the regions of Fig. 1.

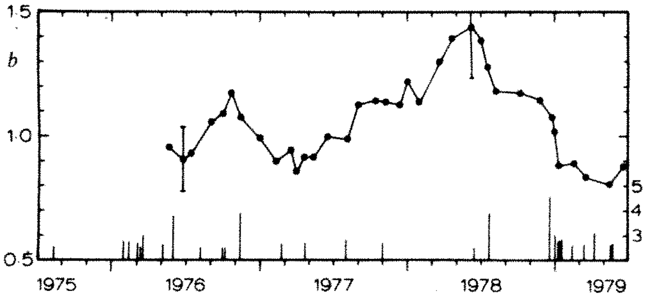


Fig. 3 b-Values for the Lake Pukaki seismic network. Magnitude threshold 1.5, window length 50 events, sliding by 10 events. Earthquakes of magnitude 3.5 and greater are also shown.

The Fiordland area (Fig. 1*D*) is very active, and if there are precursory changes in b for the large events shown, it will be necessary to decrease the search area to isolate them. Precursory changes for individual earthquakes, if they occur, are obscured by the high level of activity. The same problem is also likely to occur in Japan or any highly active area. It may be intensified in Fiordland where the source regions of earthquakes are generally small.

It is interesting to simulate these windowed b-value calculations using a synthetic data set with an exponential frequency distribution. The b-value so calculated does, of course, fluctuate. In fact the probability that the observed changes in b with real data occurred by chance is significant (if the earthquake population does actually have a random exponential distribution), but this is not the point. The proper question is: 'Are periods of high b-value followed by large earthquakes?'. The evidence presented here suggests that they are.

Note also that b-values have been calculated only as a statistical parameter of the small samples. No extrapolation to large samples is implied. The b-value is directly related to the mean magnitude, which may be a more useful parameter, but the terminology has been retained in favour of the well-known parameter b. A high b-value implies a low mean magnitude.

The *b*-value as an earthquake predictor

The evidence presented, that a period of high b-values is followed by a large event, is alternately expressed in the gap hypothesis. Mogi's[11] "seismic gap of the second kind", defined as a region deficient in moderate earthquakes before the occurrence of a large earthquake, may find expression in a high b-value. Evison's[12] suggestion of a precursory swarm followed by a gap may also be accommodated by identifying the swarm with the initial low value of b.

The b-values in each frame of Fig. 1 are calculated for an area rather larger than the precursory gap area for any of the earthquakes. This is encouraging for a surveillance scheme, if the b-value does indeed prove to be a predictor, because a routine procedure to examine the whole country for changes in b-value need not use search areas coincident with the aseismic area related to the forthcoming shocks it is endeavouring to detect. Provided that the search areas overlap the precursory area, but are not too much larger than it, the anomaly in b-values will be detected. The maximum b-value reached is not very significant, because of the mixed population it represents. What is apparently important is the temporal behaviour.

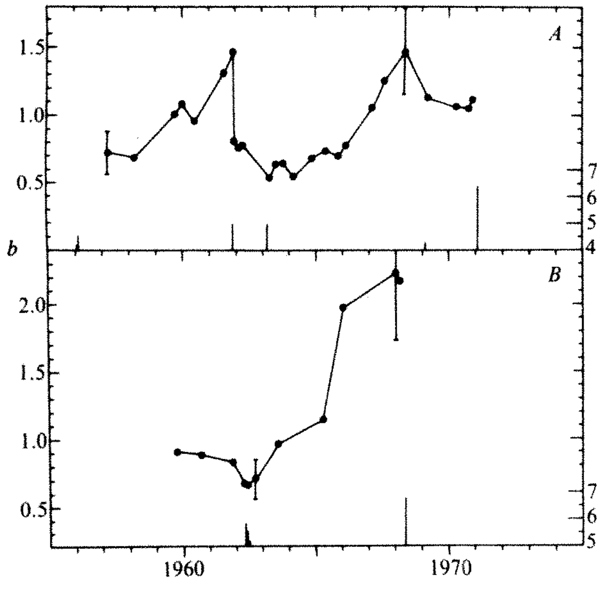


Fig. 4 b-Values for: *A*, the region of the San Fernando earthquake of 1971 (34–35°N, 118–119°W; $M \geq 2.5$) and *B*, the Inangahua earthquake of 1968 (41–43°S, 171–173°E; $M \geq 4.0$). Window length 20 events, sliding by 5 events. Earthquakes of magnitude greater than 4.0 (*A*) and 5.0 (*B*) are also shown.

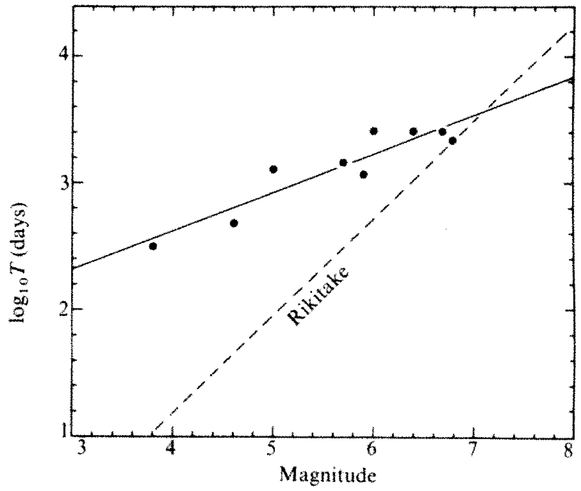


Fig. 5 Total precursor time T as a function of magnitude, for the earthquakes listed in Table 1.

Precursor time and magnitude

Total precursor times for the nine events identified are shown in Table 1, and in Fig. 5 as a function of magnitude, with the least squares regression:

$$\log T = 1.42 + 0.30M$$

For the Caracas event a magnitude of 6.7 has been used, as quoted by Rial[13], although the ISC value of m_b was only 5.7. The regression is admittedly poor, and heavily dependent on the two Pukaki earthquakes of magnitude 3.8 and 4.6. But clearly the b-value does not fall into the class of phenomena which obey Rikitake's relationship[14], also shown. The windowing technique can only reduce the observed precursor time, never increase it. This means that the ordinates plotted may be systematically small, except for the Caracas precursor which was determined from annual values. There seems to be no clear rule for the elapsed time from the b-value peak to the major earthquake, but this may be due to the windowing technique and the variety of seismicity levels in the data presented. The occurrence of foreshocks would tend to depress the b-value before a large event[15], but this is a phenomenon on a much shorter time scale than that presented here.

Conclusions

The occurrence of a period of high b-values before a large earthquake has been observed in New Zealand, California and Venezuela. Only one case is known where it may have been associated with a previous event, the occasion when high b-values followed the Inangahua earthquake of 1968, but there was also a subsequent earthquake of magnitude 5.9 with which the high b-values could be associated. There may be a decrease in b towards normal levels before the earthquake actually occurs, but it is difficult to resolve this in areas of low seismicity. The regression between magnitude and precursor time is poor, but the b-value clearly does not fall into the general class of phenomena which obey Rikitake's well known relationship.

I thank Drs E. G. C. Smith, R. Robinson and T. Hatherton for their helpful comments, and analysis staff at the Seismological Observatory for painstakingly locating many thousands of small earthquakes. Dr Smith also performed the calculations on the synthetic data sets.

Note added in proof: Ma[16] has also found evidence for high b-values before large earthquakes in China, including the Tangshan and Haicheng events.

Received 2 June; accepted 27 August 1980.

1. Gutenberg, B. & Richter, C. F. *Bull. seism. Soc. Am.* **34,** 185–188 (1944).
2. Aki, K. *Bull. Eq. Res. Inst. Univ. Tokyo* **43,** 237–238 (1965).
3. Fiedler, G. *Tectonophysics* **23,** 277–282 (1974).
4. Wyss, M. & Lee, W. H. K. *Proc. Conf. on Tectonic Problems of the San Andreas Fault System* Vol. 13, 24–42 (Stanford University, 1973).
5. Gibowicz, S. J. *Bull. seism. Soc. Am.* **63,** 517–528 (1973).
6. Gibowicz, S. J. *Tectonophysics* **23,** 283–297 (1974).
7. Robinson, R. *Phys. Earth planet. Inter.* **18,** 209–220 (1979).
8. Calhaem, I. M. *Bull. N.Z. Nat. Soc. Earthq. Engng* **12,** 7–10 (1979).
9. Hileman, J. A., Allen, C. R. & Nordquist, J. M. *Seismicity of the Southern California Region, 1 January 1931 to 31 December 1972* (California Institute of Technology, 1973).
10. Evison, F. F. *N.Z. J. Geol. Geophys.* **20,** 129–141 (1977).
11. Mogi, K. *Pure appl. Geophys.* **117,** 1172–1186 (1979).
12. Evison, F. F. *Phys. Earth planet. Inter.* **15,** P19–P23 (1977).
13. Rial, J. A. *J. geophys. Res.* **83,** 5405–5414 (1978).
14. Rikitake, T. *Bull. seism. Soc. Am.* **65,** 1133–1162 (1975).
15. Suyehiro, S. *Bull. seism. Soc. Am.* **56,** 185–200 (1966).
16. Ma, Hong-ching *Acta Geophys. Sinica* **21,** 126–141 (1978).

Actin-containing matrix associated with the plasma membrane of murine tumour and lymphoid cells

Matthew F. Mescher, Mimi J. L. Jose & Steven P. Balk

Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115

A detergent-insoluble matrix has been isolated from murine tumour cell and lymphocyte plasma membranes. The major components of this matrix include actin and four additional proteins not previously identified as cytoskeletal proteins. Labelling studies indicate that the matrix is located on the inner face of the plasma membrane. A cell-surface glycoprotein, 5′-nucleotidase, remains associated with the isolated matrix.

THE cytoplasmic face of the plasma membrane appears to be the meeting place for two sets of cellular structures intimately involved in determining the response of a cell to its environment. It is here that the transmembrane cell-surface proteins may interact with components of the cytoskeletal system. The red cell membrane has served as an extremely useful model for understanding many aspects of the plasma membrane of more complex cells[1], but whether the similarity of red cell membranes and the plasma membranes of nucleated cells extends to one of the most striking features of the red cell membrane, the 'membrane skeleton', has remained unclear. Recent studies have revealed that the red cell membrane is not composed simply of a lipid bilayer with individual proteins interacting with it. Instead, spectrin and a group of associated proteins form an extensive network at the cytoplasmic face of the bilayer and function as a membrane skeleton to stabilize the membrane. This skeletal network controls the shape and elasticity of the membrane[2,3] and the distribution of transmembrane glycoproteins[4] and intramembrane particles[5]. The network can be readily isolated by extracting membranes with the non-ionic detergent Triton X-100. Spectrin, the major skeletal protein, along with the other network proteins and some of the membrane lipids, remains insoluble[6]. The results of numerous recent studies have led to an increasingly clear picture of how these proteins are organized to form the skeletal network[7].

Spectrin is present in all mammalian and avian red cells but is not found in other cell types and a membrane skeleton similar to the spectrin network has not been demonstrated in nucleated cells. These cells have extensive filamentous cytoskeletal networks within the cytoplasm[1,8,9] and it has been proposed that there are associations between transmembrane surface proteins and components of this cytoskeletal network at the inner surface of the membrane[8,10]. These interactions may be involved in such phenomena as capping of surface molecules on lymphocytes[11,12], control of cell motility[13] and attachment to surfaces[14]. Some direct evidence for interactions between transmembrane proteins and components of the cytoskeletal system has been presented[15–17]. Whether such interactions, primarily implicated in dynamic processes such as receptor redistribution or cell motility, might also serve to stabilize the membrane is unclear. It is also unclear whether a distinct structure exists on the cytoplasmic face of the membrane, analogous to the red cell membrane skeleton. The results reported here demonstrate that a detergent-insoluble component can be isolated from the plasma membranes of murine tumour and lymphoid cells which may function as a membrane skeletal network. The network is apparently located on the inner surface of the membrane and has five major protein components. Only one of these proteins, actin, has previously been shown to be a component of cytoskeletal structures.

Plasma membrane isolation and Nonidet P-40 solubilization

Plasma membranes were isolated as described previously for P815 ascites cells[18] with only minor modifications (see Fig. 1 legend). This procedure results in a 40–50-fold purification, as assessed by 5′-nucleotidase activity, and a yield of 40–60%. Co-purification of the cell-surface H-2 antigens with the plasma membrane fraction (ref. 18 and unpublished data) confirmed the results obtained by assaying 5′-nucleotidase. Membranes have been purified from P815, EL4, RDM-4 and CH1 murine tumour cells and normal murine spleen cells with essentially identical results. Examination of the purified membranes by electron microscopy shows them to be in the form of vesicles of relatively heterogeneous size, ~0.1–1.2 μm in diameter[18]. Small amounts of endoplasmic reticulum are also apparent, as indicated by the presence of some membrane vesicles with bound ribosomes. Based on yields obtained, the plasma membrane accounts for 1–3% of the total cell protein, in good agreement with values reported by others[19].

0028-0836/81/030139—06$01.00

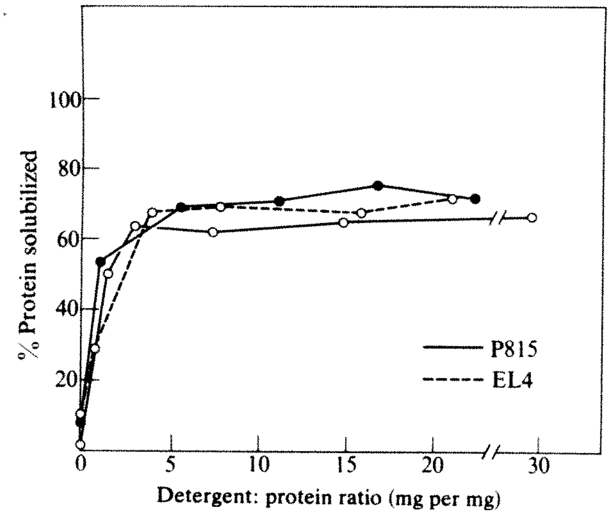


Fig. 1 Solubilization of P815 and EL4 plasma membrane protein with Nonidet P40. Plasma membranes were purified as described previously with minor modifications[18]. The cells were lysed by nitrogen cavitation in Earle's salts buffered at *p*H 7.2 with 10 mM sodium phosphate (PBES) containing 1 mM PMSF and 10 mM iodoacetamide. Following differential centrifugation, the membrane pellet containing predominantly plasma membrane and endoplasmic reticulum was dispersed in buffer by brief sonication in a bath sonicator (10 × 10 s in a Branson model B-12) and loaded on a sucrose step gradient. The purified plasma membrane fraction was resuspended in PBES at 2–5 mg ml^{-1} and stored frozen (−20 °C) until use. Detergent solubilization was done by resuspending membranes in PBES with vigorous vortexing and adding 1% NP-40 in PBES to give the desired detergent to protein ratio. Protein concentration was held constant at 0.5 mg ml^{-1}. The samples were incubated on ice for 20 min with frequent vortexing and then centrifuged for 45 min at 100,000*g*. After centrifugation, the supernatants were removed and the insoluble pellets resuspended in PBES. Protein content of the supernatant and pellets was assayed by the method of Lowry *et al.*[33] in the presence of 1% SDS using bovine serum albumin as the standard. Values shown are averages of duplicate determinations. -●—●- and -○—○-, results obtained using two separate P815 membrane preparations. --○--○--, EL4 membrane solubilization.

Treatment of lymphocyte plasma membranes with the non-ionic detergent Nonidet P-40 (NP-40) solubilizes much of the membrane protein, including histocompatibility antigens and surface immunoglobulins, but leaves an insoluble residue which can be pelleted by centrifugation at 100,000*g*. When solubilization of P815 (a mastocytoma) and EL4 (a thymoma) membranes was examined in more detail, it was found that the amount of protein solubilized increased with increasing detergent up to a detergent to protein ratio of approximately 4 : 1, which gave 70–80% solubilization (Fig. 1). The remaining 20–30% of the protein remained insoluble at detergent to protein ratios of up to 30 : 1. The amount of protein remaining was constant above a detergent to protein ratio of 4 : 1 regardless of the concentration of detergent (from 2–20 mg ml^{-1}). Re-extraction of the insoluble residue with NP-40 resulted in no additional solubilization. Identical results were obtained using Triton X-100 (TX-100) in these experiments and those described below. The detergent-insoluble material formed a tight, translucent pellet on centrifugation at 100,000*g*. In comparison, the unextracted membranes formed a white, opaque pellet.

Protein composition of the detergent-insoluble fraction

Relative to red cell membranes, the plasma membranes of nucleated cells contain a larger number of proteins. At least 50 membrane polypeptide chains can be resolved by SDS-polyacrylamide gel electrophoresis of P815 plasma membranes (Fig. 2*a*, lane 1). Examination of the NP-40-soluble and -insoluble fractions of P815 membranes revealed that most of the proteins were completely solubilized while a small number remained selectively insoluble (Fig. 2). The major components of the insoluble fraction were polypeptides having apparent molecular weights (MWs) of 70,000 (70K), 69,000 (69K), 42,000 (42K), 38,000 (38K), and 36,000 (36K). The p70, p69, p38 and p36 proteins were found only in the detergent-insoluble fraction. Proteins at 42K were present in both the soluble and insoluble fractions. The occurrence of the minor bands seen in the detergent-insoluble fraction, particularly the low-molecular weight proteins near the dye front, varies between membrane preparations. Some of these may be products of proteolysis, although phenylmethylsulphonyl fluoride (PMSF, 1 mM) and iodoacetamide (10 mM) are present during membrane purification and solubilization.

Comparison of membranes and whole cells labelled with ^{125}I by lactoperoxidase-catalysed iodination suggested that most of the detergent-insoluble proteins are not exposed in intact cells. p70, p69, p38 and p36 are prominently labelled proteins on isolated membranes (Fig. 2*b*, lane 4). In contrast, they are not detected when proteins from labelled intact cells, containing comparable amounts of radioactivity, are examined (Fig. 2*b*, lane 5). p42 is also a major labelled protein on membranes but proteins in the same molecular weight range are labelled on intact cells; therefore, these results do not exclude the possibility that some of the p42 is present at the cell surface. The efficient labelling of p70, p69, p38 and p36 by lactoperoxidase on

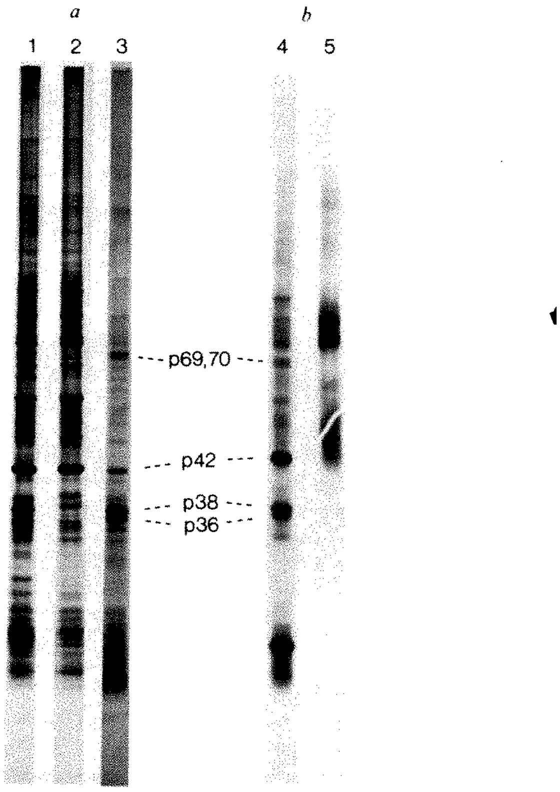


Fig. 2 *a*, The protein composition of membranes and detergent-soluble and -insoluble fractions. P815 tumour cell membranes were prepared as described in the Fig. 1 legend. Membranes were solubilized in 0.5% NP-40 (also as described in Fig. 1 legend) at a detergent to protein ratio of 5 : 1 (w/w). The samples were then analysed by SDS-polyacrylamide slab gel electrophoresis following reduction with 5% β-mercaptoethanol. Electrophoresis was done using the buffer system of Laemmli[34] on a running gel consisting of a 5–15% polyacrylamide gradient. Gels were stained with Coomassie Brilliant blue following electrophoresis. Lane 1, intact membranes; lane 2, proteins solubilized by 0.5% NP-40 (100,000*g* supernatant); lane 3, detergent-insoluble fraction (100,000*g* pellet). *b*, ^{125}I-labelled membranes and intact cells. Cells were labelled by lactoperoxidase-catalysed iodination as previously described[35]. After labelling and washing to remove free iodine, cells were solublized in SDS gel application buffer, reduced with 5% β-mercaptoethanol and electrophoresed as described above. Isolated membranes were labelled by the same procedure, washed and prepared for electrophoresis in the same way. Following electrophoresis, the gel was dried and ^{125}I-labelled proteins visualized by autoradiography. Lane 4, ^{125}I-labelled membranes; lane 5, ^{125}I-labelled whole cells. Stained samples (lanes 1–3) and iodinated samples (lanes 4, 5) were run on separate SDS gels. MWs are based on standard proteins run on each gel.

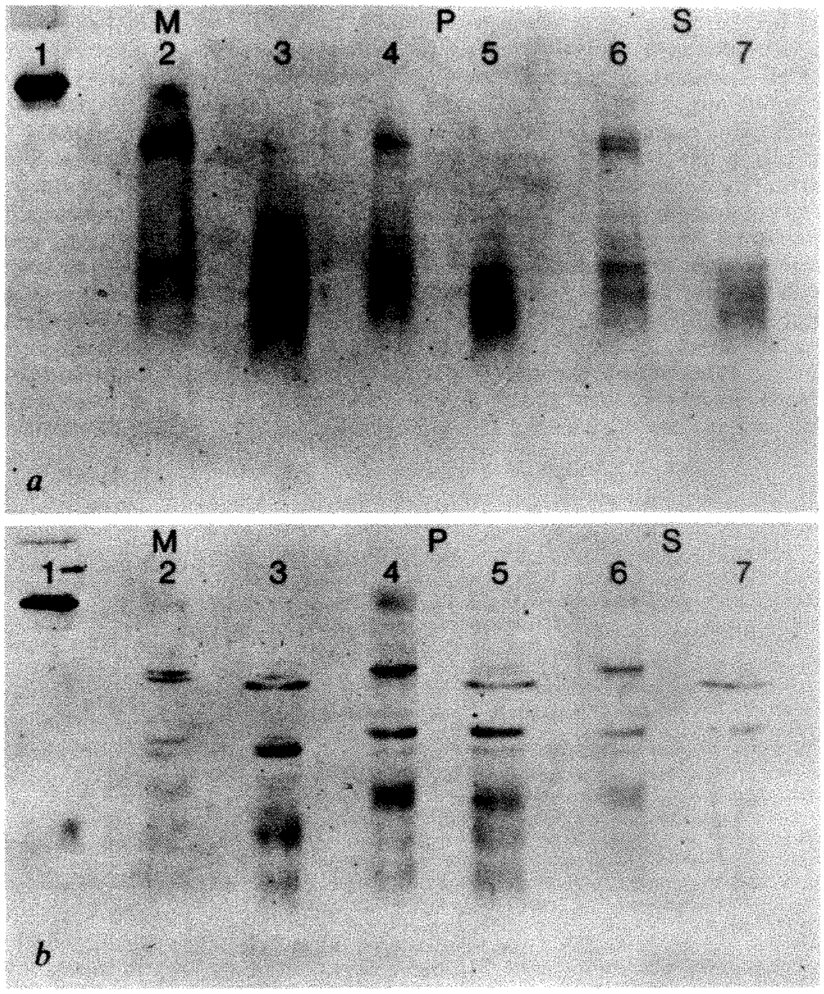


Fig. 3 Limited proteolysis peptide mapping of p42 from the NP-40-soluble and -insoluble fractions of P815 tumour cell membranes. Mapping was done by the procedure of Cleveland *et al.*[22]. Plasma membranes were solubilized with NP-40, the soluble and insoluble fractions electrophoresed on 7.5% polyacrylamide gels and the gels lightly stained. The p42 bands of each fraction were sliced out and soaked for 30 min in a solution containing 0.125 M Tris-HCl, *p*H 6.8, 0.1% SDS and 1 mM EDTA. The slices were then placed in the sample wells of a 15% polyacrylamide gel and the same buffer containing 10% glycerol and a given amount of protease layered on top. Electrophoresis was then done in the usual manner with the exception that the current was turned off for 30 min just before the bromophenol blue tracking dye entered the running gel. Gels were stained for protein with Coomassie Brilliant blue. Proteolytic enzymes used were: *a*, *Streptomyces* protease, type VI (Sigma); *b*, V8 protease (Miles) from *Staphylococcus aureus*. Protein samples for mapping were: M, rabbit muscle actin with: no protease (lane 1), 0.05 μg protease (lane 2) and 0.25 μg protease (lane 3). P, p42 from the detergent-insoluble fraction (100,000*g* pellet) with 0.05 μg (lane 4) and 0.25 μg (lane 5) protease. S, p42 from the detergent-soluble fraction (100,000*g* supernatant) with 0.05 (lane 6) and 0.25 μg (lane 7) protease.

isolated membranes and not on intact cells suggests that these proteins are located on the inner face of the membrane.

Actin has been demonstrated to be present in plasma membranes from a variety of cell types[20], including lymphocytes[21], and the 42K MW protein present in P815 membranes co-migrates with rabbit skeletal muscle actin. When the p42s of the NP-40-soluble and -insoluble fractions were compared by limited proteolysis peptide mapping[22], it was found that both gave peptide patterns which confirmed their identity as actin (Fig. 3). Fragments obtained using *Streptomyces* protease (Fig. 3*a*) were very similar for rabbit muscle actin and p42 and those obtained using V8 protease (Fig. 3*b*) showed only small differences between the two. Small differences in peptide patterns are expected for the same protein from different species. No significant differences with either protease could be seen for the soluble and insoluble forms of p42. It thus seems that P815 membranes have two forms of actin, one soluble in NP-40 and the other making up a major component of the detergent-insoluble fraction. As mentioned above, re-treating the insoluble fraction with detergent results in no further solubilization of the actin.

Results of limited proteolysis peptide mapping of p38 and p36 ruled out the possibility that one or both of these proteins might be actin degradation products (results not shown). Although p70, p69, p38 and p36 have not been identified, we should point out that they do not have the same molecular weights as any of the proteins which have been shown to interact with actin or to be components of cytoskeletal structures. α-actinin, 100,000 MW, has been suggested to be involved in the attachment of microfilaments to the plasma membrane[23,24] and Hoessli *et al.*[25] have shown that murine lymphocytes synthesize a protein having the expected properties of α-actinin. A protein of 100,000 MW is present in intact membranes from P815 cells but is completely solubilized by NP-40 (Fig. 2).

A detergent-insoluble component comprising the same major proteins has been found in membranes of all the tumour cell lines examined and in normal murine lymphocytes. Such cells include EL4 (a thymoma), RDM-4 (a lymphoma), CH1 (a B-cell lymphoma) and normal murine spleen lymphocytes. In all cases, NP-40 solubilization at detergent to protein ratios of >4:1 leaves 20–30% of the protein insoluble and the major protein components of the insoluble fraction have MWs of 70K, 69K, 42K, 38K and 36K.

The cell-surface enzyme 5′-nucleotidase[26,27] was also found to be a component of the detergent-insoluble fraction. When P815 membranes were solubilized as shown in Fig. 1, 90–98% of the 5′-nucleotidase activity was recovered in the detergent-insoluble fraction at all detergent to protein ratios tested. This was also true for EL4, CH1 and normal spleen cell membranes. 5′-Nucleotidase purified from pig lymphocytes has been reported to have a MW of 130,000 (ref. 28). It is unlikely, therefore, that any of the major proteins seen in the insoluble fraction is 5′-nucleotidase; the enzyme is probably present in too low an amount to be visualized by Coomassie blue staining.

5′-nucleotidase is present on a wide variety of cell types including most lymphocytes, and cytochemical studies have demonstrated that it is localized predominantly on the cell surface[26,27]. Experiments were done to confirm the surface location of the 5′-nucleotidase activity of P815 cells. The cells were incubated for 1 h at 4 °C in the presence or absence of 5 mg ml^{-1} concanavalin A (Con A), an inhibitor of 5′-nucleotidase. They were then washed to remove unbound lectin, lysed by freeze–thaw, centrifuged at 100,000*g* to pellet membranes and the 5′-nucleotidase activity of the pellets measured. Cells which had been exposed to Con A had about 70% less activity than untreated cells. Control experiments demonstrated that free Con A, which could cause inhibition of internal 5′-nucleotidase, was not available at the time of cell lysis. Following washing to remove unbound Con A, treated and untreated cells were mixed and lysed by freeze–thaw. The 5′-nucleotidase activity of the membranes was the sum of that obtained for each type of cell lysed separately, indicating that no free Con A was available to inhibit the activity of the untreated cell membranes. Similarly, when both types of membrane were mixed and assayed, no further inhibition was obtained. These results indicate that at least 70% of the membrane-bound 5′-nucleotidase activity of P815 cells is located at the cell surface. This probably represents a minimum estimate as cell aggregation during the incubation with Con A may result in some of the surface enzyme becoming inaccessible to the lectin.

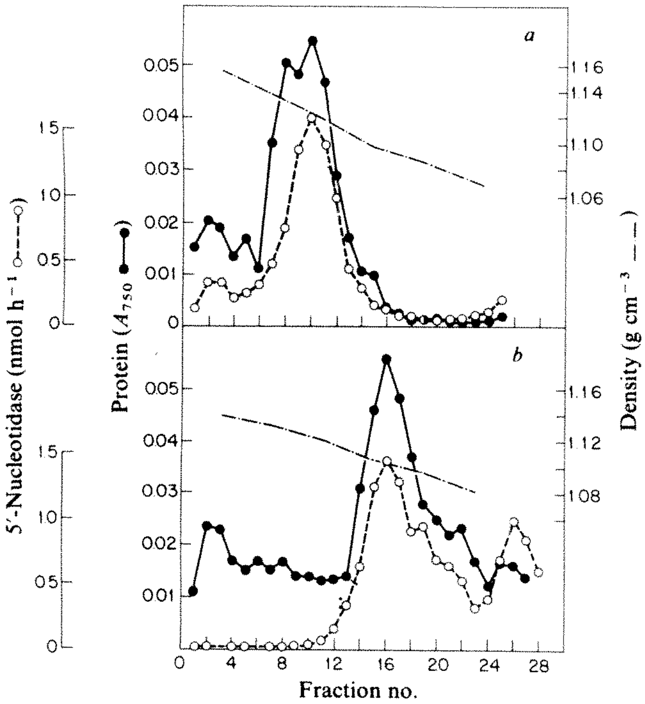


Fig. 4 Sucrose density gradient profiles of P815 plasma membranes and the detergent-insoluble fraction. Detergent-insoluble material was prepared as described in Fig. 1 legend using Triton X-100 in PBES at a detergent to protein ratio of 5:1 and washed once in PBES. Membranes or detergent-insoluble material were resuspended in PBES at a concentration of 2 mg ml^{-1} and layered on top of a 20–40% sucrose gradient (in 10 mM Tris-HCl, pH 7.2) with a 50% cushion at the bottom. The gradients were centrifuged for 18 h at 18,000 r.p.m. in a Beckman SW-27 rotor. Fractions were then collected from the bottom of the tubes and assayed for protein (A_{750}) and 5′-nucleotidase activity[16]. *a*, Density gradient profile obtained with 500 μg of P815 membranes; *b*, profile obtained with 400 μg of detergent insoluble material.

Density gradient centrifugation of the detergent-insoluble fraction

The occurrence of a specific set of proteins, including part of the membrane actin, in the NP-40-insoluble fraction suggested that this material might consist of a membrane skeletal network structure and not simply membrane proteins insoluble in NP-40 in aqueous buffer. We therefore examined the insoluble fraction by equilibrium density gradient centrifugation to determine if it behaved as a discrete structure and as an integral part of the membrane and not a contaminant in the membrane preparation.

P815 plasma membranes were found at a density of 1.125 ± 0.001 g cm^{-3} (average of three experiments ± s.d.) on overnight centrifugation in a 20–40% sucrose gradient (see Fig. 4). When the TX-100-insoluble fraction from P815 membranes was examined in the same way, the protein and 5′-nucleotidase peak had shifted to a density of 1.107 ± 0.005 g cm^{-3} (average of three experiments ± s.d.). Most (>70%) of the 5′-nucleotidase activity remained associated with the protein peak. The TX-100-insoluble fraction has a somewhat lower density than intact membranes. Experiments using ^{3}H-labelled TX-100 demonstrated that the lower density was not due to residual detergent bound to the protein. Less than 0.02 mg of detergent per mg of protein could be detected in the washed, detergent-insoluble material. Some of the membrane lipid remains bound to the red cell cytoskeleton on isolation using TX-100[6] and this is also the case for the detergent-insoluble material from P815 cells. This fraction contains 1.0–1.4 μmol phospholipid per mg of protein (chloroform/methanol extractable organic phosphate). Preliminary examination of this lipid by TLC indicates that some of the membrane lipids remain selectively associated with the insoluble fraction. Studies are in progress to identify these lipids.

The protein composition of fractions across the TX-100-insoluble peak from a density gradient was examined by SDS-gel electrophoresis (not shown). No indication of any separation of one protein species from another could be detected. Examination of fractions across the intact membrane peak from a density gradient in the same way also gave no indication of any separation of the major insoluble proteins from other membrane components. These results strongly suggest that the proteins of the detergent-insoluble fraction are components of the plasma membrane and form a structure which remains intact after solubilization of the other membrane proteins and much of the lipid. Examination of the insoluble fraction by electron microscopy further supports this suggestion.

The detergent-insoluble plasma membrane matrix

The results described above strongly suggested that plasma membranes have a matrix of protein at their cytoplasmic face which can be isolated on the basis of its insolubility in non-ionic detergent. The detergent-insoluble preparations were examined by electron microscopy to determine if their appearance was consistent with this conclusion. Plasma membranes from P815 cells isolated by the procedure used here are predominantly in the form of closed vesicles[18]. They are heterogeneous in size, ranging from ~0.1 to 1.2 μm in diameter, and have a typical unit membrane appearance. No filamentous material associated with the membranes is apparent in the preparations[18]. The detergent-insoluble fraction from these membranes also seems to be largely in the form of closed structures (Fig. 5) also having a heterogeneous size distribution. A unit membrane structure does not seem to be present.

The appearance of the detergent-insoluble fraction in thin section supports the suggestion that this material consists of a skeletal matrix structure isolated from the inner surface of the membrane and not simply of aggregates of insoluble protein. No protein layer adjacent to the bilayer is apparent in thin sections of intact membranes, suggesting that the skeletal network is very closely associated with the lipid bilayer. The spectrin network of red blood cells is similarly not apparent in thin sections of intact red cell membranes[6].

Further evidence for the presence of a matrix associated with the plasma membrane was provided by experiments done with whole cells. Treatment of intact lymphoid cells with NP-40 (or TX-100) lyses the cells and solubilizes the membrane and cytoplasmic proteins while leaving the nuclei intact. This procedure is frequently used for solubilizing cell-surface antigens before immunoprecipitation. If the detergent-insoluble proteins form a continuous matrix on the inner surface of the plasma membrane, it might be expected that this matrix would be trapped around the nucleus and sedimented by low speed centrifugation along with the intact nuclei following cell solubilization. The 5′-nucleotidase activity associated with the detergent-insoluble matrix provides a convenient marker for determining if this is the case. When P815 cells were solubilized with 0.5% NP-40 (at 5 × 10^{6} cells ml^{-1}) and centrifuged for

10 min at 1,000g, ~60% of the 5′-nucleotidase activity was found in the nuclear pellet. In these same conditions, >95% of ^{125}I-labelled cell-surface proteins were found in the supernatant. Of the 40% of 5′-nucleotidase present in the low speed supernatant, approximately half (19%) could be sedimented by centrifugation at 100,000g. When the same experiment was done using normal murine spleen cells, essentially identical results were obtained.

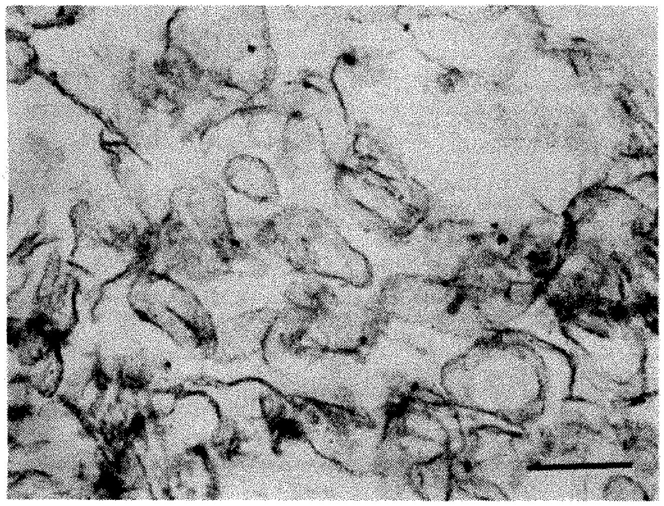


Fig. 5 Electron micrograph of thin section of the detergent-insoluble fraction (100,000g pellet from P815 plasma membrane). Scale bar, 0.2 μm.

Discussion

The results described here strongly suggest that actin and a small, discrete set of additional proteins form a matrix associated with the plasma membrane of murine tumour and lymphoid cells. Labelling studies using lactoperoxidase indicate that the matrix is probably localized at the inner face of the membrane, although further studies will be necessary to confirm this. The matrix is in some respects similar to the spectrin matrix of red blood cells in that it remains selectively insoluble in non-ionic detergents and some of the membrane lipid remains associated with it. Not surprisingly, the protein components are different from those of the spectrin matrix, with the exception of actin.

Unlike erythrocytes, nucleated cells have extensive internal membrane systems and it is therefore more difficult to obtain pure plasma membranes. Although the plasma membranes used in these experiments are highly purified, it is clear that some contaminants are present, including small amounts of endoplasmic reticulum and some small, electron-dense material[18]. Several of the observations reported here argue strongly, however, that the detergent-insoluble matrix is derived from the plasma membrane and not from contaminants in the preparation. The detergent-insoluble matrix proteins account for a relatively large fraction (~20%) of the total protein in the preparations and these proteins co-purify with the plasma membranes on density gradient centrifugation. Although some contaminants are apparent in electron micrographs of the purified membranes, none has the appearance of the isolated matrix. The large size and closed structure of the isolated matrix make it very likely that they are derived from the plasma membrane vesicles. The association of 5′-nucleotidase with the detergent-insoluble matrix also provides strong evidence for the plasma membrane origin of this material. The majority of this enzyme is localized on the plasma membrane of P815 (see results) and other cells[26,27].

The functional role of the plasma membrane matrix described here is not known. It may have a structural role, serving to stabilize the membrane bilayer. The location of the matrix also makes these proteins likely candidates for mediating interactions between transmembrane proteins and the filament systems of the cytoskeleton. The matrix proteins may serve as anchorage points on the inner membrane surface for components of the cytoskeletal system.

In other work being done in this laboratory[29] it was found that reconstituted vesicles containing lipid, purified H–2 antigen and the isolated matrix proteins could be formed by mixing the components in deoxycholate-containing buffer and dialysing to remove the detergent. The H–2 antigen present in these vesicles was much more effective in stimulating generation of an *in vitro* cytolytic T-lymphocyte response than vesicles having only H–2 and lipid. Whether this greater effectiveness of the vesicles containing matrix proteins results from interaction of the H–2 with the matrix is now being investigated.

Evidence for the interaction of at least one cell-surface glycoprotein with the matrix is provided by the observation that 5′-nucleotidase remains associated with the detergent-insoluble fraction. This association may occur through interaction of the enzyme with either a matrix protein(s) or with the lipid which remains associated with the matrix. It is difficult to speculate about the functional role for this association, as the role of the cell-surface 5′-nucleotidase activity is not well understood. It is interesting, however, that the cell-surface 5′-nucleotidase displays an unusual property with respect to inhibition by Con A which may be related to its association with the matrix. Carraway *et al.*[30] have shown that the 5′-nucleotidase activity of MAT-A and MAT-C1 cells is inhibited in a non-cooperative manner by Con A. Cooperative inhibition was observed, however, if the cells were first treated with colchicine, cytochalasin B or D, or dibucaine. No direct correlation could be found between the induction of cooperativity and changes in Con A receptor redistribution or cell morphology.

The association of 5′-nucleotidase with the matrix and the resulting lack of solubilization when whole cells are lysed with NP-40 suggest a potential problem in attempting to identify cell-surface proteins by immunoprecipitation. Solubilization with non-ionic detergent, centrifugation and immunoprecipitation from the soluble supernatant, has become a standard procedure for determining the cell-surface proteins recognized by antisera and for studying cell-surface antigens. A cell-surface antigen which interacts with the matrix as the 5′-nucleotidase does would be difficult, if not impossible, to isolate by this procedure.

Actin occurs in two forms in the plasma membranes examined here, one soluble in detergent and the other associated with the insoluble matrix. Gruenstein *et al.*[31] have previously shown that membranes from 3T3 mouse fibroblasts and HeLa cells have two forms of actin. Treatment of the membranes in conditions which favour actin filament depolymerization resulted in removal of 80% (3T3) and 60% (HeLa) of the actin while the rest remained associated with the membrane. This actin fraction may be a component of a membrane matrix similar to that described here. We have noted that the soluble and insoluble forms of actin described here usually have a slight difference in mobility on SDS gels and we are now investigating the possibility that the two forms differ structurally.

With the exception of actin, the protein components of the matrix do not have molecular weights consistent with their being previously described components of the cytoskeletal network. Moore *et al.*[32] have examined the TX-100-insoluble fraction from sarcoma 180 ascites cell membranes. The membranes were isolated from cells which had been treated with Zn^{2+}, a procedure which seems to stabilize the interaction of cytoskeletal elements with the membrane. The TX-100-insoluble fraction of these membranes included proteins which the authors tentatively identified as actin-binding protein, α-actinin and actin. Several other polypeptides were also present in these preparations. Our failure to observe components of the cytoskeletal network associated with the membrane matrix may be due to failure to preserve these associations during membrane isolation and/or solubilization. It seems that actin and the other components observed in our studies are sufficient to form a stable, detergent-insoluble matrix.

Preliminary experiments have suggested that human peripheral lymphocytes have a plasma membrane matrix similar to that found for murine cells. The major protein components had the same mobilities on SDS gels as those found for the matrix components of murine cells. It will be of interest to determine whether a matrix is associated with the plasma membranes of other cell types (for example fibroblasts) which have more extensive cytoskeletal networks than lymphocytes and, if so, whether the same major protein components constitute the matrix.

This work was supported by National Institute of Allergy and Infectious Diseases grant 1-PO1-AI-14732. We thank Dr Morris Karnovsky for the electron micrographs and for helpful discussions during this work.

Received 1 July; accepted 4 November 1980.

1. Bretscher, M. S. & Raff, M. C. *Nature* **258,** 43–49 (1975).
2. LaCelle, P. L. & Kirkpatrick, F. H. in *Erythrocyte Structure and Function* (ed. Brewer, G. J.) 535–557 (Liss, New York, 1975).
3. Sheetz, M. P. & Singer, S. J. *J. Cell Biol.* **73,** 638–646 (1977).
4. Nicolson, G. L. & Painter, R. G. *J. Cell Biol.* **59,** 395–406 (1973).
5. Elgsaeter, A. & Branton, D. *J. Cell Biol.* **63,** 1018–1036 (1974).
6. Yu, J., Fischman, D. A. & Steck, T. L. *J. supramolec. Struct.* **1,** 233–248 (1973).
7. Lux, S. E. *Nature* **281,** 426–429 (1979).
8. Nicolson, G. L. *Biochim. biophys. Acta* **457,** 57–108 (1976).
9. Lazarides, E. *Nature* **283,** 249–256 (1980).
10. Edelman, G. M. *Science* **192,** 218–226 (1976).
11. dePetris, S. & Raff, M. C. *Eur. J. Immun.* **2,** 523–535 (1972).
12. Schreiner, G. F., Fujiwara, K., Pollard, T. D. & Unanue, E. R. *J. exp. Med.* **145,** 1393–1398 (1977).
13. Korn, E. D. in *Cold Spring Harb. Conf. Cell Proliferation* Vol. 13 (eds Goldman, P., Pollard, T. & Rosenbaum, J.) 623–630 (Cold Spring Harbor Laboratory, New York, 1967).
14. Rees, D. A., Lloyd, C. W. & Thom, D. *Nature* **267,** 1240128 (1977).
15. Koch, G. & Smith, M. *Nature* **273,** 274–278 (1978).
16. Flanagan, J. & Koch, G. *Nature* **273,** 278–281 (1978).
17. Carter, W. R. & Hakomori, S. *J. biol. Chem.* **253,** 2867–2874 (1978).
18. Lemonnier, F., Mescher, M., Sherman, L. & Burakoff, S. *J. Immun.* **120,** 1114–1120 (1978).
19. Crumpton, M. J. & Snary, D. *Contemp. Topics molec. Immun.* **3,** 27–56 (1974).
20. Pollard, T. D. & Weihing, R. R. *CRC crit. Rev. Biochem.* **2,** 1–65 (1964).
21. Barber, B. H. & Crumpton, M. J. *FEBS Lett.* **66,** 215–220 (1976).
22. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252,** 1102–1106 (1977).
23. Bourguignon, L. Y. W. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5031–5035 (1977).
24. Loor, F. *Nature* **264,** 272–273 (1976).
25. Hoessli, D., Rungger-Brändle, E., Jockusch, B. M. & Gabbiani, G. *J. Cell Biol.* **84,** 305–314 (1980).
26. Misra, D. N., Gill, T. J. III & Estes, L. W. *Biochim. biophys. Acta* **352,** 455–461 (1974).
27. Uusitalo, R. J. & Karnovsky, J. J. *J. Histochem. Cytochem.* **25,** 87–96 (1976).
28. Dornand, J., Bonnafous, J.-C. & Maini, J.-C. *Eur. J. Biochem.* **87,** 459–465 (1978).
29. Herrmann, S. H. & Mescher, M. F. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
30. Carraway, K. L., Doss, R. C., Huggins, J. W., Chesnut, R. W. & Carraway, C. A. C. *J. Cell Biol.* **83,** 529–543 (1979).
31. Gruenstein, E., Rich, A. & Weihing, R. R. *J. Cell Biol.* **64,** 223–234 (1975).
32. Moore, P. B., Ownby, C. L. & Carraway, K. L. *Expl Cell Res.* **115,** 331–342 (1978).
33. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193,** 265–275 (1951).
34. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
35. Mescher, M. F. & Pollock, R. R. *J. Immun.* **123,** 1155–1161 (1979).
36. Avruch, J. & Wallach, D. F. H. *Biochim. biophys. Acta* **233,** 334–347 (1971).

Intrachromosomal gene conversion in yeast

Hannah L. Klein & Thomas D. Petes

Department of Microbiology, The University of Chicago, 920 East 58th street, Chicago, Illinois 60637

We have shown that the yeast Saccharomyces cerevisiae *has a mechanism by which information from one gene can be transferred non-reciprocally to a repeated copy of the gene on the same chromosome. This intrachromosomal gene conversion may be important in maintaining sequence homogeneity within families of repeated eukaryotic genes.*

GENETIC exchange is usually thought to be an interchromosomal event occurring between homologous chromosomes in a diploid cell. This exchange of genetic information between two chromosomes may be either a reciprocal or a non-reciprocal (gene-conversion) event. A second type of genetic exchange is intrachromosomal recombination (or conversion). This term has been used to describe both recombination between chromatids attached to the same centromere (sister-strand recombination) and recombination between repeated genes within a single chromatid (intrachromatid recombination). Experimental evidence for both types of intrachromosomal recombination has been obtained. Using a selectable marker inserted into the tandemly repeated ribosomal DNA genes of yeast, T.D.P.[1] and Szostak and Wu[2] have shown that unequal sister chromatid exchange occurs in meiosis and mitosis. Sister chromatid exchange has also been observed in *Drosophila*[3] and in cultured mammalian cells[4]. Reciprocal recombination between repeated elements within a single DNA strand (intrachromatid recombination) has been seen within the 2-μm plasmid of yeast[5–9], as a rare recombination event of yeast chromosome III associated with a mating-type switch[10] and as a mitotic excision of a duplicated *LEU*2 gene from chromosome III (ref. 11). No clear example of non-reciprocal intrachromosomal exchange has been reported although mating-type switching in yeast can be interpreted as a unidirectional gene-conversion event[12,13].

The technique of yeast transformation provides an opportunity for examining the more general occurrence of intrachromosomal gene exchange. Hinnen *et al.*[11] showed that a transforming recombinant plasmid containing an insertion of yeast DNA usually integrates into the yeast chromosome by a homologous recombination event. As shown in Fig. 1 (after Fig. 2 of Hinnen *et al.*[11]), the result of this integration is a chromosome that contains a non-tandem duplication of a portion of yeast genome with the duplicated segments separated by bacterial plasmid sequences. In the experiments of Hinnen *et al.*[11], a recombinant plasmid containing the yeast *LEU*2$^+$ gene was integrated into a chromosome containing a *leu*2 mutation on chromosome III. We have shown that the mutation in the mutant copy can be transferred to the wild-type *LEU*2 gene by an *in vivo* meiotic intrachromosomal gene conversion. We do not know whether the transfer of information occurs as the result of an intrachromatid interaction or an unequal association between sister chromatids. Whatever the details of the interaction, the existence of such a mechanism may have important consequences for the evolution of repeated genes.

Isolation of a strain containing a duplicated *LEU*2 gene

The haploid yeast strain AH22[11] contains two frameshift mutations, *leu*2-3 and *leu*2-112, in the *LEU*2 gene on chromosome

III. This strain was transformed with the plasmid CV9 (obtained from J. Hicks). The CV9 plasmid contains a wild-type *LEU*2 yeast gene on a *Pst* restriction fragment that is inserted into the *Pst* site of pBR322[14]. Forty-nine independent Leu$^+$ transformants were obtained and one of these, AH22/CV9-12, was analysed in detail.

In the first experiment, AH22/CV9-12 was crossed to a haploid of opposite mating type (strain 2262; ref. 1) that carried a *leu*2 mutation. When the resulting diploid was sporulated and tetrad analysis performed[15], 41 of 42 tetrads examined segregated $2^+{:}2^-$ for the Leu2 phenotype. This segregation pattern indicates that the *LEU*2$^+$ transforming DNA had integrated at a single site in the host genome[11]. To determine the site of integration, we crossed AH22/CV9-12 to a *LEU*2$^+$ haploid strain HK-3A to form the diploid HK-12. Tetrad analysis of HK-12 showed segregation patterns of 4 Leu$^+$ to 0 leu$^-$ spores in more than 90% of the tetrads. This result indicates that the transforming *LEU*2$^+$ gene is closely linked to the normal *LEU*2 locus on chromosome III. The presumed arrangement of the *LEU*2 genes in AH22/CV9-12 is shown in Fig. 1 and the arrangement of markers in the diploid HK-12 (after meiotic DNA synthesis) is given in Fig. 2*a*. Additional evidence confirming these arrangements will be given below.

Of 306 tetrads analysed from the diploid HK-12, 294 tetrads segregated 4 Leu$^+$ to 0 leu$^-$ spores and 12 segregated 3 *LEU*$^+$ to 1 *leu*$^-$ spores. As described below, a detailed genetic analysis of tetrads showing the aberrant $3^+{:}1^-$ pattern provided evidence for intrachromosomal gene conversion.

Genetic analysis of aberrant tetrads

The diploid strain HK-12 had the genotype:

$$\frac{\text{a}\ \ leu2\text{-}3\ \ leu2\text{-}112\ \text{plus CV9 insertion}}{\alpha \qquad\qquad LEU2^+}$$

$$\frac{his4}{HIS4^+}\ \ \frac{TRP1^+}{trp1}\ \ \frac{ADE6^+}{ade6}\ \ \frac{MET13^+}{met13}\ \ \frac{TYR7^+}{tyr7}.$$

For purposes of our analysis, the most important genetic markers other than the *LEU*2 genes were the *HIS*4 gene, located on chromosome III centromere-distal to *LEU*2, and *TRP*1, a tightly centromere-linked marker on chromosome IV. These markers were used to monitor gene–gene or gene–centromere recombination on chromosome III by standard tetrad analysis[15].

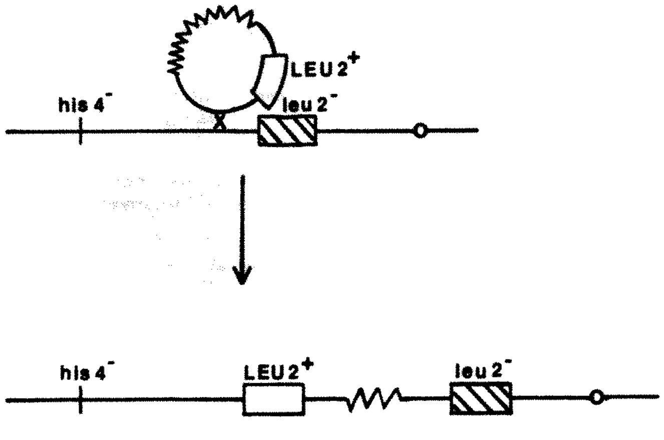


Fig. 1 Derivation of AH22/CV9-12 (after Hinnen *et al.*[11]). The recombinant plasmid CV9, consisting of a *Pst* fragment carrying the yeast *LEU*2$^+$ gene inserted into pBR322, is shown to integrate into the yeast chromosome III during transformation by homologous recombination. The exchange event is shown here to occur in flanking sequences that are centromere-distal to the *LEU*2 coding region. The resulting transformant contains duplicated *LEU*2 genes separated by pBR322 sequences. The thin line represents yeast sequences, the open boxes the yeast *LEU*2$^+$ gene, the lined boxes the yeast *leu*2$^-$ gene and the wavy line pBR322 sequences.

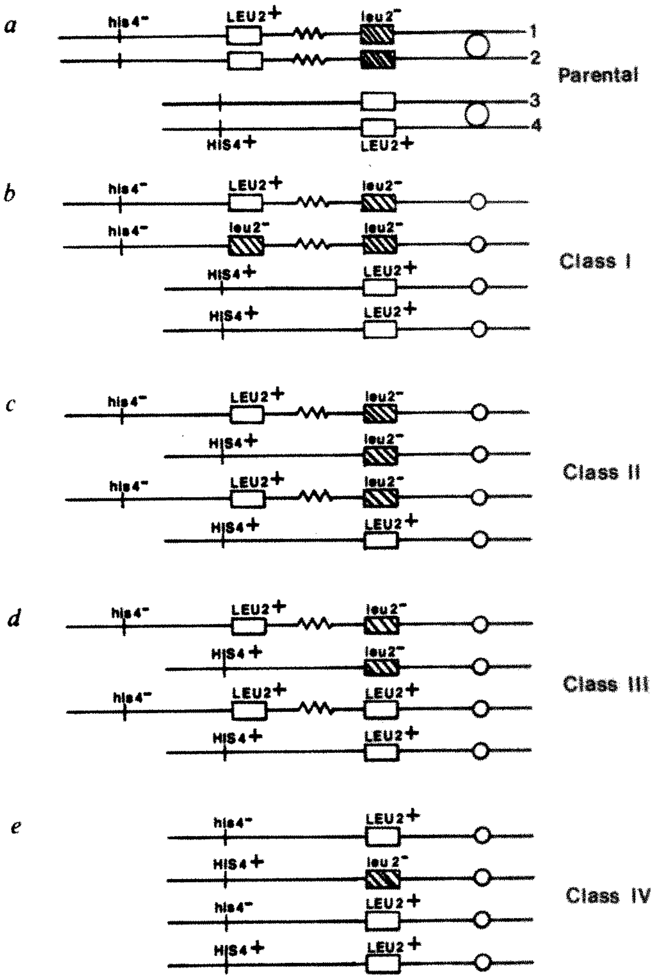


Fig. 2 Arrangement of *LEU*2 genes on chromosome III following transformation. The arrangement of *LEU*2 genes and the *HIS*4 gene on chromosome III in the diploid HK-12 is shown. Also shown are the order of these genes in tetrads from classes I, II, III and IV. *a* Shows the parental configuration of chromosome III in HK-12 following premeiotic DNA synthesis. The chromatids are labelled 1–4. *b* Shows an example of the chromosome III configuration in a class I tetrad. In some tetrads, reciprocal recombination events along chromosome III (noted in Table 1) have also occurred. *c* Shows an example of the chromosome III configuration in the class II tetrads. *d* Shows the chromosome III configuration in the class III tetrad. *e* Shows the chromosome III configuration in the class IV tetrad. The thin line represents yeast sequences, the open boxes the *LEU*2$^+$ gene, the lined boxes the *leu*2$^-$ gene and the wavy line pBR322 sequences.

Each spore clone from the aberrant tetrads was analysed in several ways. First, each strain was checked by standard procedures for auxotrophic requirements. Second, we examined each strain for the presence or absence of pBR322 sequences by using the colony hybridization technique[11]. Third, DNA from each strain was examined using the Southern blotting technique[16] to determine whether the spore had one or two copies of the *LEU*2 gene per chromosome. The details of the Southern analysis are described in Fig. 3 legend.

The procedures described above were done with all the spores from the aberrant tetrads. For some spores, additional tests were done. For those spores that were phenotypically leu$^-$, we did standard allelism tests[17] to determine whether they contained the *leu*2-3 allele, the *leu*2-112 allele, both mutant alleles or a new mutation in the *LEU*2 gene.

Some of the spores derived from the aberrant tetrads contained two copies of the *LEU*2 gene and were phenotypically Leu$^+$. In such spores, there could be either two *LEU*2$^+$ genes or one *LEU*2$^+$ gene and one mutant *leu*2 gene. To distinguish between these possibilities, we did a mutability test of the spores. The rationale for the test is that it should be more difficult to isolate a *leu*2-3 or *leu*2-112 mutation from a strain that has two wild-type *LEU2* genes than from a strain that has one wild-type *LEU2* gene and one *leu2*-3 *leu2*-112 mutant gene. As will be described in more detail below, by comparing the frequency of induced *leu2*-3 and *leu2*-113 mutations with the frequency of other *leu*$^-$ mutations, we could distinguish between two copies of the wild-type *LEU2* gene and one copy each of the wild-type and mutant *LEU2* genes.

Table 1 summarizes the genetic analysis of the aberrant tetrads of HK-12. On the basis of these data, we concluded that there were several mechanisms by which aberrant tetrads could be produced: class I (intrachromosomal gene conversion), class II (interchromosomal gene conversion), class III (reciprocal interchromosomal exchange) and class IV (multiple genetic events).

The four classes of aberrant tetrads

Class I tetrads were the most common aberrant segregation; 6 of the 12 aberrant tetrads were in this group. As described below, our analysis indicates that this class is the result of intrachromosomal gene conversion. The diagnostic feature of class I tetrads is the presence of two mutant copies of the *LEU2* gene (and associated pBR322 plasmid) in the leu$^-$ mutant spore (Table 1, Figs 2*b*, 3). Thus, a chromosome which originally had one mutant *leu2* gene and one wild-type *LEU2* gene has been changed to one which has two mutant genes. Allelism tests showed that the *leu2* mutations present in the leu$^-$ spores are derived from the *leu2* mutation of the untransformed parent AH22 and do not result from new *leu2*$^-$ alleles (Table 1). Both mutant alleles *leu2*-3 and *leu2*-112 (class IA) or one mutant allele *leu2*-3 (class IB) occur in the converted *leu2* gene. The mutational changes during intrachromosomal gene conversion do not involve deletions of any significant size, as no differences in the size of *LEU2*-specific restriction fragments are observed by Southern analysis (Fig. 3). Also, the *leu2*$^-$ mutants obtained in class IB can be reverted to the Leu$^+$ phenotype (H.K., data not shown).

The tetrads of HK-12 that have not undergone a conversion event (Fig. 2*a*) should have four wild-type *LEU2* genes and two mutant *leu2* genes within the four haploid spores. If class I tetrads represent a gene conversion event of a wild-type *LEU2* gene to a mutant *leu2* gene, they would be expected to have three *LEU*$^+$ genes and three mutant *leu2* genes. As previously

Table 1 Genetic analysis and segregation of aberrant tetrads

Class	Tetrad	*LEU2**	pBR322	*HIS4*	*TRP1*	No. and deduced phenotype of *LEU2* genes†	Homologous recombination on chromosome III
IA	IA	+	−	+	−	1 (+)	
	B	+	−	+	−	1 (+)	
	C	*leu2*-3, 2-112	+	−	+	2 (−−)	
	D	+	+	−	+	2 (+−)	
	2A	+	−	+	−	1 (+)	
	B	+	+	−	+	2 (+−)	Exchange between *HIS4* and pBR322 sequences
	C	*leu2*-3, 2-112	+	+	+	2 (−−)	
	D	+	−	−	−	1 (+)	
	3A	+	−	+	+	1 (+)	
	B	+	+	−	−	2 (+−)	Exchange centromere-proximal to pBR322 sequences
	C	*leu2*-2, 2-112	+	−	+	2 (−−)	
	D	+	−	+	−	1 (+)	
IB	4A	*leu2*-3	+	−	−	2 (−−)	
	B	+	−	+	+	1 (+)	
	C	+	−	+	+	1 (+)	
	D	+	+	−	−	2 (+−)	
	5A	+	−	+	−	1 (+)	
	B	+	+	−	+	2 (+−)	
	C	*leu2*-3	+	−	+	2 (−−)	
	D	+	−	+	−	1 (+)	
	6A	+	+	−	−	2 (+−)	
	B	+	−	+	+	1 (+)	Exchange centromere-proximal to pBR322 sequences
	C	*leu2*-3	+	−	+	2 (−−)	
	D	+	−	+	−	1 (+)	
II	7A	+	−	+	+	1 (+)	
	B	+	+	−	−	2 (+−)	Exchange between *HIS4* and pBR322 sequences
	C	+	+	+	−	2 (+−)	
	D	*leu2*-3, 2-112	−	−	+	1 (−)	
	8A	+	+	−	+	2 (+−)	
	B	*leu2*-3, 2-112	−	+	−	1 (−)	Exchange centromere-proximal pBR322 sequences
	C	+	−	+	+	1 (+)	
	D	+	+	−	−	2 (+−)	
	9A	+	+	−	+	2 (+−)	
	B	+	+	−	−	2 (+−)	Exchange centromere-proximal to pBR322 sequences
	C	*leu2*-3, 2-112	−	+	−	1 (−)	
	D	+	−	+	+	1 (+)	
	10A	+	+	−	+	2 (+−)	
	B	+	−	+	+	1 (+)	Exchange centromere-proximal pBR322 sequences
	C	*leu2*-3, 2-112	−	+	−	1 (−)	
	D	+	+	−	−	2 (+−)	
III	11A	+	+	−	+	2 (++)	
	B	+	−	+	+	1 (+)	Exchange centromere-proximal pBR322 sequences
	C	+	+	−	−	2 (+−)	
	D	*leu2*-3, 2-112	−	+	−	1 (−)	
IV	12A	+	−	−	−	1 (+)	
	B	+	−	+	+	1 (+)	Exchange centromere-proximal to *HIS4*
	C	*leu2*-112	−	+	−	1 (−)	
	D	+	−	−	+	1 (+)	

* The *leu2* alleles present in the leu$^-$ spores were determined by standard allelism tests. These involved mating each leu$^-$ spore with *leu2*-3 and *leu2*-112 testers of the appropriate mating type. Following growth of the diploids, the crosses were irradiated with UV rays, grown overnight in the dark and replica-plated to synthetic medium lacking leucine. The appearance of papillations due to mitotic recombination induced by the UV-ray treatment was taken as an indication of non-allelism. When only one *leu2* allele was observed in the leu$^-$ spore, it is assumed that the second copy of the *LEU2* gene contains the original double mutation *leu2*-3 *leu2*-112, although it is possible that the *leu2*-112 allele may have been lost by intrachromosomal gene conversion.

† This column shows the phenotype of the *LEU2* genes present in each spore. Minus signifies a mutant *leu2* gene. Plus signifies a wild-type *LEU2* gene. This interpretation is based on the data shown in Table 2.

discussed, the only ambiguous spore class in terms of the number of wild-type *LEU*2 alleles is that in which the spore is phenotypically Leu$^+$ and has two copies of the *LEU*2 gene. Such spores could contain either two wild-type *LEU*2 alleles or one mutant and one wild-type allele. As discussed previously, these two alternatives should be distinguished by testing the frequency with which *leu*2-3 and *leu*2-112 mutations can be recovered. To do this test, we grew a suspension of the cells being tested in rich medium. The cells were then treated with UV-ray irradiation (75% survival) and grown up overnight in the dark. The culture was then plated for single colonies on rich medium and replica-plated to synthetic medium lacking leucine. Complementation tests and allelism tests were done on the leucine-requiring mutants that were isolated.

The data for spores of class I as well as the original parental strains are shown in Table 2. Two results indicate that all the tested spores from class I tetrads contain one wild-type *LEU*2 gene and one mutant *leu*2 gene. First, in all cases it was easy to recover mutations at the *LEU*2 locus. For each class I strain tested, mutations at the *LEU*2 locus accounted for more than half of the total *leu*$^-$ mutations. In addition, the spectrum of mutant alleles recovered from the class I spores was similar to those recovered from the parental strain AH22/CV9-12. Second, at least half of the *leu*2 mutations produced by UV light are *leu*2-3, *leu*2-112 or both *leu*2-3 and *leu*2-112. We assume that the UV-ray treatment is not specifically inducing mutations at these loci, but is 'uncovering' the pre-existing mutations. The effect of the radiation is likely to be a stimulation of excision of the wild-type *LEU*2 gene or stimulation of gene conversion. The mutant *leu*2 genes remaining in the spore after UV-ray treatment are unlikely to be deletion mutants because several of the mutants are revertible (H.K., unpublished data). Also, the *leu*2-3 *leu*2-112 mutants derived from spore HK-121D show recombination with a different *leu*2 allele (*leu*2-1).

In summary, the mutability analysis indicates that the class I Leu$^+$ spores that contain two *LEU*2 genes have one wild-type gene and one mutant gene. Within the class I tetrads, therefore, there are a total of three wild-type *LEU*2 genes and three mutant *leu*2 genes (Fig. 2*b*). These results strongly support intrachromosomal gene conversion as the mechanism that generates class I tetrads. We do not know whether the intrachromosomal conversion event occurs as the result of an interaction within one chromatid or an unequal interaction between sister chromatids. For either type of interaction, we assume that the conversion event involves heteroduplex formation followed by mismatch repair[18], as for classical gene conversion.

Class II tetrads result from interchromosomal gene-conversion events. They are distinguished from class I tetrads by the segregation pattern of pBR322 sequences with the leu$^-$ spore. In class I tetrads the leu$^-$ spore contains pBR322 sequences, whereas the class II leu$^-$ spore does not (Table 1). Analysis of the types of leu$^-$ clones generated by UV-ray irradiation of the Leu$^+$ spores containing pBR322 sequences in the class II tetrads shows that both of these spores in each tetrad contain a mutant *leu*2 gene in addition to the wild-type *LEU*2 gene (Table 2). Thus, the class II tetrads have three mutant *leu*2 genes and three wild-type *LEU*2 genes. In contrast to the class I tetrads, the mutant *leu*2 genes are distributed over three chromosomes rather than two (Fig. 2*c*). This pattern of segregation requires the same interchromosomal interaction between homologues as occurs in 'classical' gene conversions[15].

Class III represents a reciprocal recombination event; only one such tetrad was found. In this tetrad, the mutability test of the spore HK12-11A (which is Leu$^+$ and has two copies of the *LEU*2 genes) yields a different result from similar tests of class I and II spores. The low frequency of *leu*2 mutations in this strain (Table 2) suggests that the spore HK12-11A has two wild-type *LEU*2 genes. Within the four spores, therefore, this class III tetrad has two mutant *leu*2 genes and four wild-type *LEU*2 genes (as shown in Fig. 2*d*). The simplest explanation of this segregation pattern is that a reciprocal recombination event occurred between the homologous chromosomes (chromatids 2 and 3 in Fig. 2*a*). This event occurred in the yeast sequences flanked by pBR322 and the mutant *leu*2 gene on strand 2 and homologous sequences centromere-distal to the wild-type *LEU*2 gene on strand 3 (see also Table 1).

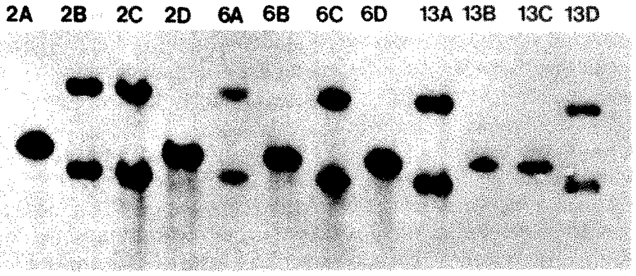


Fig. 3 Quantification of *LEU*2 genes. An autoradiogram of two class I tetrads and one control tetrad is shown. Nuclear DNA was isolated[29] from class I tetrads 2 and 6 and a control tetrad called tetrad 13 that showed 4$^+$: 0$^-$ segregation for the *LEU*2 gene. DNA was digested with *Bam*HI endonuclease. There is no *Bam*HI site within the yeast DNA inserted into plasmid pBR322 in the CV9 vector. The pBR322 plasmid contains one *Bam*HI site. Thus, a yeast strain with one *LEU*2 gene should have one *Bam*HI fragment containing the *LEU*2 gene. A yeast strain carrying two *LEU*2 genes separated by pBR322 sequences will have two *Bam*HI fragments, each containing one *LEU*2 gene. Following digestion of nuclear DNA with *Bam*HI, the fragments were separated by agarose gel electrophoresis on a horizontal 0.7% agarose slab gel run at 1 V cm^{-1} for 36 h. DNA was transferred to nitrocellulose[16]. The filters were hybridized to a DNA probe labelled with ^{32}P by nick translation[30]. The probe used was the large *Eco*RI fragment of the CV9 recombinant plasmid. The fragment contains approximately one-half of the yeast fragment inserted into pBR322 and half of the pBR322 plasmid.

The segregation pattern of this tetrad provides preliminary evidence that the orientation of genes in the AH22/CV9-12 parent is (*his*4)-(*LEU*2$^+$ gene)-(pBR322 sequences)-(*leu*2 mutant gene)-(centromere). If the order of the wild-type *LEU*2 gene and mutant gene were reversed, one should not obtain a spore with one mutant *leu*2 gene and no pBR322 sequences as the result of a single recombination event.

The most difficult segregation pattern (class IV) to explain is that found for tetrad 12. In this tetrad, no pBR322 sequences are present in any of the spores. In addition, only four copies of the *LEU*2 gene are present instead of the expected six copies. As both of the pBR322 sequences are missing from the tetrad, it is possible that two independent meiotic excisions of the plasmid

Table 2 Number of UV ray-induced leu$^-$ segregants

Strain	*leu*2-3 2-112	*leu*2-3	*leu*2-112	*leu*2 mutant alleles different from *leu*2-3 and *leu*2-112	*leu*$^-$ complementation groups other than *LEU*2
1D	18	0	9	5	13
2B	15	10	7	6	21
3B	3	2	0	2	1
4D	18	2	0	3	12
5B	7	0	2	1	4
6A	2	0	6	6	3
7B	11	1	4	3	3
7C	9	3	2	1	2
8A	10	0	0	0	0
8D	0	3	1	10	2
9A	3	0	1	1	2
9B	2	0	0	3	1
10A	1	2	1	2	3
10D	4	1	1	2	4
11A	0	0	0	0	10
11C	16	0	3	2	9
AH22/CV9-12	12	5	7	6	0
HK-3A	0	0	0	0	2

Selected spores from Table 1 that were Leu$^+$, but contained two copies of the *LEU*2 gene were grown overnight in rich medium. Approximately 10^6 cells were irradiated with UV rays to give 75% survival. Cells were grown overnight in the dark, plated for single colonies on rich medium and replica-plated to synthetic medium lacking leucine. Leucine-requiring mutants were picked and retested for the leucine requirement. The *leu*2 alleles were determined by the allelism test used in Table 1.

occurred (one associated with a conversion). Alternatively, if the excision of the plasmid was associated with heteroduplex formation between the duplicated *LEU*2 genes during G_1 of meiosis, replication of the resulting chromosome could produce a class IV tetrad.

Our discussion of classes I–IV aberrant tetrads has emphasized the number and types of *LEU*2 genes in each spore. By examining the segregation of the *HIS*4 and *TRP*1 genes, we could also determine the position of other recombination events along chromosome III (between *HIS*4 and *LEU*2 and *LEU*2 and the centromere). This information is also summarized in Table 1.

In conclusion, a non-tandem duplication consisting of a mutant *leu*2 gene and a wild-type *LEU*2 gene can be altered by a number of genetic mechanisms. The most common alteration is an intrachromosomal gene conversion. Interchromosomal gene-conversion events and other alterations are also found. As will be discussed below, we believe intrachromosomal gene conversion may have important consequences in the evolution of repeated genes.

Discussion

Repeated cycles of unequal sister chromatid exchange have been suggested as a mechanism for maintaining homogeneity among highly repeated sequences such as the ribosomal DNA sequences, satellite DNA, histone and the immunoglobulin genes[3,19,20]. Recently, this mechanism has been extended to the globin genes following the observation that the restriction enzyme maps of the duplicated α-globin genes in humans and primates are highly conserved within a species[21,22]. Although unequal sister-strand recombination undoubtedly occurs in some systems[1,2,22–24], it is not clear that such exchanges are the mechanism by which sequence homogeneity is maintained. There are at least two inherent disadvantages in maintaining sequence homogeneity among repeated genes by unequal sister-strand exchange. First, unequal sister-strand exchange involves a change in gene dosage. For some classes of repeated genes, therefore, one would expect selection against the products of the exchange. Second, if the repeated genes on a chromosome are not arranged in tandem, but are separated by unique DNA, unequal sister chromatid exchange will result in the deletion of these unique sequences from one chromatid.

We believe that certain experimental results can be most easily explained by an intrachromosomal event. For example, in certain individuals, the γ-globin genes contain a restriction endonuclease pattern which is different from that found in most humans[25]. Both copies of the γ gene on a single chromosome contain this change. Jeffreys[25] has proposed that some type of intrachromosomal recombination event maintains this polymorphism. DNA sequencing of γ-globin genes suggests the possibility of a gene-conversion event[26].

Although intrachromosomal gene conversion has the advantage that a mutation can be duplicated or deleted in a single generation without changing gene dosage, it is still possible, of course, that unequal sister-strand recombination is also important in maintaining sequence homogeneity, as these events can be quite frequent[1,2]. In the present study six tetrads contained an intrachromosomal gene conversion and there were no examples of an unequal sister-strand exchange. We do not know whether the lack of unequal exchanges is the result of the small sample size, the small size of the duplicated segment or some other factor. We suggest that in some systems, unequal sister-strand exchange may be more important as a mechanism for changing gene dosage than as a mechanism for maintaining sequence homogeneity. The relative importance of intrachromosomal gene conversion and unequal sister-strand exchange in maintaining sequence homogeneity may depend on many factors, for example, whether the duplicated genes are in tandem or separated by unique DNA sequences, the number of repeated units in the family or the DNA sequence within the repeated unit.

In the data reported here, intrachromosomal gene conversion, resulting in two mutant *leu*2 genes, was observed in 6/306 unselected tetrads, a frequency of 2%. If the conversion of a mutant *leu*2 allele to a wild-type *LEU*2 allele is assumed to occur at the same rate, the overall frequency of intrachromosomal gene conversion is estimated to be 4%. This frequency is clearly much higher than the frequency of mutation. Consequently, this mechanism should be suffcient to maintain the homogeneity of certain classes of repeated genes (such as those genes that code for ribosomal RNA) within the yeast genome. Note also that for some classes of repeated genes, intrachromosomal gene conversion may not function often enough to remove all sequence diversity. The conversion frequencies may depend on such factors as the specific base sequence, the size of the repeated gene or the distance between repeated units.

Finally, there is the question of whether the intrachromosomal gene conversion that we have observed occurs naturally in yeast or is induced by the insertion of the recombinant plasmid CV9. Although we cannot entirely rule out the latter possibility, we believe that it is unlikely. Using different recombinant plasmids and different selectable yeast genes, Jackson and Fink (personal communication) and Falco and Botstein (personal communication) have found similar results. In the experiments of Jackson and Fink, both meiotic and mitotic conversion events were observed. In addition, Munz and Leupold[27], in reversion studies of suppressor mutations in the yeast *Schizosaccharomyces pombe*, have obtained evidence that is consistent with the occurrence of interchromosomal gene-conversion events among repeated tRNA genes. Scherer and Davis[28] have recently demonstrated that similar non-homologous interchromosomal gene-conversion events occur in *Saccharomyces cerevisiae*.

We thank Dorothy Stamenkovitch and Tuula Vuorio for technical assistance, Drs R. Esposito, S. Klapholtz, S. Lindquist and R. Farber for comments and criticism concerning the manuscript, and Drs Fink and Botstein for communicating to us unpublished results. This work was supported by USPHS training grant GM 07543-02 (HLK) and USPHS grant CA 19265-04A1 (HLK). T.D.P. was also supported by NIH grant GM 24110, March of Dimes grant 5-207 and Research Career Development Award 1KO4 AG00077.

Received 6 August; accepted 29 October 1980.

1. Petes, T. D. *Cell* **19,** 765–774 (1980).
2. Szostak, J. W. & Wu, R. *Nature* **84,** 426–430 (1980).
3. Tartof, K. D. *Cold Spring Harb. Symp. quant. Biol.* **38,** 491–500 (1973).
4. Latt, S. A. & Schreck, R. R. *Am. J. hum. Genet.* **32,** 297–313 (1980).
5. Beggs, J. D., Guerineau, M. & Atkins, J. F. *Molec. gen. Genet.* **148,** 287–294 (1976).
6. Guerineau, M., Grandchamp, C. & Slonimski, P. P. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3030–3034 (1976).
7. Hollenberg, C. P., Degelman, A., Kustermann-Kuhn, B. & Roger, H. D. *Proc natn. Acad. Sci. U.S.A.* **73,** 2072–2076 (1976).
8. Gubbins, E. J., Newlon, C. S., Kann, M. D. & Donelson, J. E. *Gene* **1,** 185–207 (1977).
9. Livingston, D. M. & Klein, H. L. *J. Bact.* **129,** 472–481 (1977).
10. Strathern, J., Newlon, C., Herskowitz, I. & Hicks, J. *Cell* **18,** 309–319 (1979).
11. Hinnen, A., Hicks, J. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1929–1933 (1978).
12. Hicks. J., Strathern, J. & Herskowitz, I. *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 457–462 (Cold Spring Harbor Laboratory, New York, 1977).
13. Klar, A. J. S. & Fogel, S. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4539–4543 (1979).
14. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
15. Mortimer, R. K. & Hawthorne, D. C. *The Yeasts* Vol. I (eds Rose, A. M. & Harrison, J. S.) 286–460 (Academic, London, 1969).
16. Southern, E. *J. molec. Biol.* **98,** 503–518 (1975).
17. Roman, H. & Jacob, F. *Cold Spring Harb. Symp. quant. Biol.* **23,** 155–160 (1958).
18. Meselson, M. S. & Radding, C. N. *Proc. natn. Acad. Sci. U.S.A.* **72,** 358–361 (1975).
19. Smith, G. P. *Cold Spring Harb. Symp. quant. Biol.* **38,** 507–513 (1973).
20. Hood, L., Campbell, J. H. & Elgin, S. C. R. *A. Rev. Genet.* **9,** 305–353 (1975).
21. Lauer, J., Shen, C. -K. J. & Maniatis, T. *Cell* **20,** 119–130 (1980).
22. Zimmer, E. A., Martin, S. K., Beverly, S. M., Kan, Y. W. & Wilson, A. C. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2158–2162 (1980).
23. Goossens, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 518–521 (1980).
24. Manning, R. F. & Gage, L. P. *J. biol. Chem.* (in the press).
25. Jeffreys, A. J. *Cell* **18,** 1–10 (1979).
26. Slightom, J. L., Blechl, A. E. & Smithies, O. *Cell* **21,** 627–638 (1980).
27. Munz, P. & Leupold, U. *Alfred Benzon Symp.* **16** (in the press).
28. Scherer, S. & Davis, R. W. *Science* **209,** 1380–1384 (1980).
29. Williamson, D. H. & Fennel, D. J. *Meth. Cell Biol.* **12,** 335–351 (1975).
30. Schachat, F. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **38,** 371–381 (1973).

Ordering of mouse immunoglobulin heavy chain genes by molecular cloning

Akira Shimizu, Naoki Takahashi, Yuriko Yamawaki-Kataoka, Yasuyoshi Nishida, Tohru Kataoka & Tasuku Honjo

Department of Genetics, Osaka University Medical School, Kita-ku, Osaka 530, Japan

We have determined the order of the mouse immunoglobulin $\gamma 1$, $\gamma 2b$, $\gamma 2a$ and ε genes by molecular cloning of overlapping chromosomal segments. The results clearly demonstrate that the order is 5′–$\gamma 1$–(21 kilobases)–$\gamma 2b$–(15 kilobases)–$\gamma 2a$–(15 kilobases)–ε–3′. There seem to be no J regions at the 5′ side of each constant region gene so far obtained except for the μ gene and these constant region genes seem to have repetitive sequences characteristic of switch (S) regions at their 5′ side.

IMMUNOGLOBULIN heavy (H) chain genes comprise a family of variable region (V) genes and several constant region (C) genes which are classified, in mouse, into five major classes: μ, γ, α, δ and ε. Genetic studies using allotype markers have clearly demonstrated that the V_H and C_H genes of a single chromosome are coordinately expressed—this is called *cis* expression[1-3]. During differentiation, a given lymphocyte seems to be able to associate a single V region sequentially with two or more different classes of H chain C region[4-11]. In this H chain class switch, the direction of the switch is always from μ to γ or α.

Using cDNA–mRNA hybridization kinetics, we have shown that specific C_H genes are deleted in mouse myelomas depending on which C_H genes are expressed[12], and have proposed that a chromosomal segment which intervenes recombining V_H and C_H genes is deleted to bring the V_H and C_H genes close together. Based on the deletion profiles of C_H genes in various myeloma tumours producing different C_H genes, a linear arrangement of C_H genes has been proposed as 5′–μ–$\gamma 3$–$\gamma 1$–$\gamma 2b$–$\gamma 2a$–α–3′. Subsequently, more direct evidence has been presented that deletion of the intervening DNA segment accompanies the H chain class switch[13-18]. Such studies using Southern blot[19] hybridization experiments also support the proposed order of C_H genes.

Comparative studies of rearranged and germ-line H chain genes have led to proposals[20-26] for a molecular mechanism of H chain class switch. According to this model, a complete H chain gene is formed by at least two types of the recombination event. The first type of recombination takes place between a given V_H, a J_H and a D gene segment, completing a V region gene[23-25]. The second type of recombination is required to switch a C_H gene and occurs between S_μ and S_γ (or S_α) regions which are located at the 5′ flanking regions of respective C_H genes. This model postulates two important structural features in the H chain gene organization: (1) the presence of only one set of J_H region genes in the 5′ side to the μ gene and (2) the presence of the S region before each C_H gene.

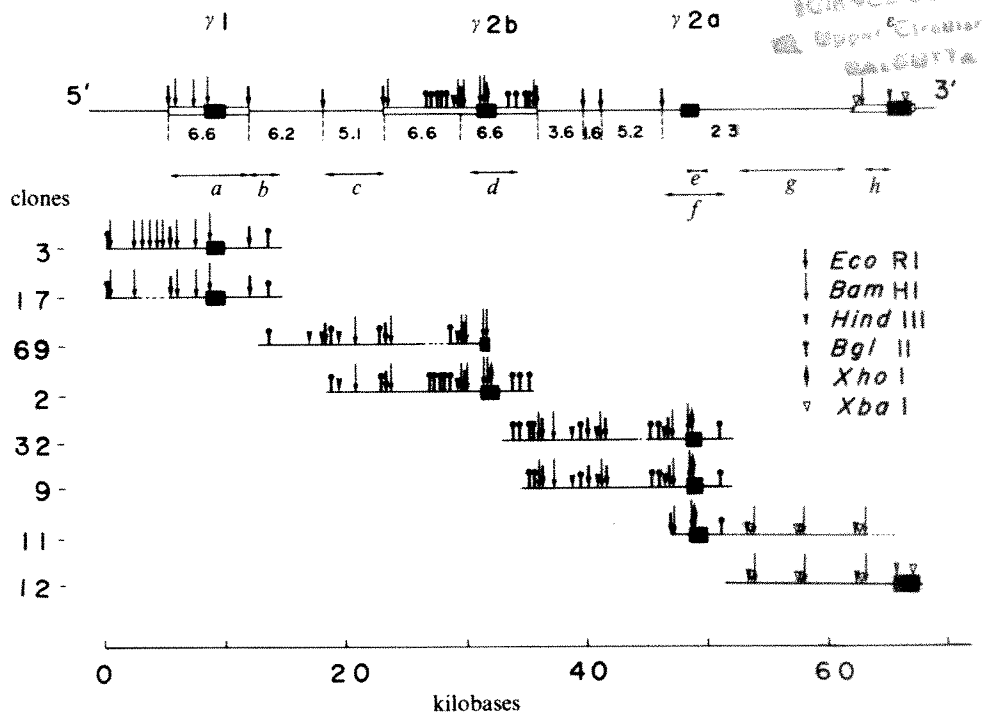


Fig. 1 Restriction endonuclease cleavage maps of overlapping cloned fragments between $\gamma 1$ and ε genes. We used a recombinant phage library[42] for cloning DNA segments that contain immunoglobulin genes or their flanking regions, except for clone 11. The Charon 4A phage[43] library of BALB/c mouse embryo DNA (provided by Dr P. Leder) contained partial *Hae*III digests of embryonic mouse DNA with average length 18 kilobases. These fragments were ligated with *Eco*RI linkers and incorporated into phage DNA. The 20-kilobase insert of clone 11 was partially purified from *Eco*RI digests of RPC5 myeloma DNA by RPC5 column chromatography and agarose gel electrophoresis as described previously[27]. The partially purified DNA fragments were ligated with the outer fragments of Charon 4A phage DNA. Screening was done on 230×230 mm tryptone broth plates containing 5–20×10⁴ plaques by an *in situ* hybridization method[44]. Appropriate restriction endonuclease fragments of immunoglobulin genes were labelled with ^{32}P by nick translation[45] to a specific activity of 30–200 c.p.m. per pg and used as hybridization probes. Nitrocellulose filters were washed as described elsewhere[17]. A plaque containing an immunoglobulin gene or its flanking region was picked up and the phages were grown in liquid cultures to prepare DNA as described[46]. At the top is a schematic representation of the chromosomal segment containing the $\gamma 1$, $\gamma 2b$, $\gamma 2a$ and ε genes. Open bars indicate DNA segments cloned previously: the $\gamma 1$ gene (IgH2)[27], the $\gamma 2b$ gene (IgH22, Igγ2b-26)[28,30,33] and the ε gene (Igε – 1)[34]. Structural genes are shown as solid boxes with direction of transcription from left to right, determined by nucleotide sequence determination. Numbers under the top line indicate sizes of *Eco*RI fragments in the germ-line gene. Horizontal arrows indicate fragments used as probes for screening or Southern hybridization. The insert of each clone is shown by a horizontal line with restriction sites. Clones were aligned by restriction endonuclease cleavage mapping and Southern hybridization analysis (see text and Figs 2, 4). Restriction digests were examined by electrophoresis in 0.5–0.7% agarose gels. It is possible that very small fragments (for example, 100 base pairs) may have been missed. *Bgl*II and *Xho*I sites at the 3′ side of the $\gamma 2a$ gene and *Xba*I sites at the 5′ side of the $\gamma 2a$ gene were not determined. Broken lines in clones 11, 17, 32 and 69 indicate deletions.

0028-0836/81/030149—05$01.00

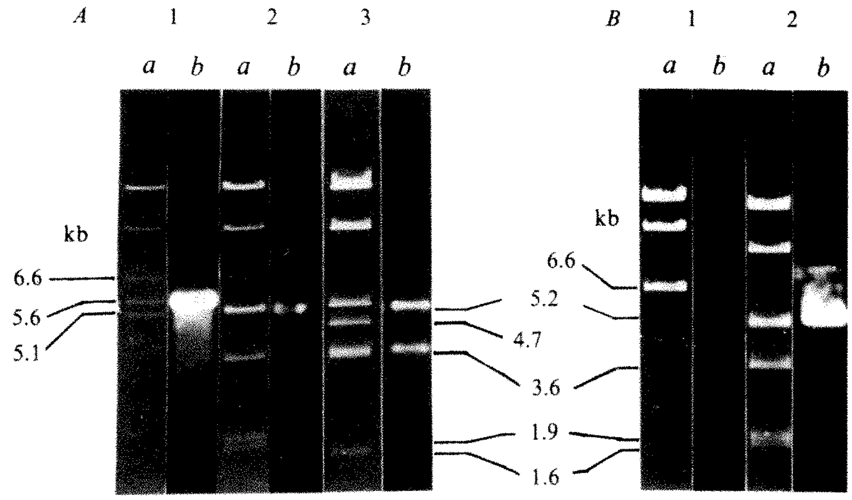


Fig. 2 Southern hybridization of isolated clones linking between the γ2b and γ2a genes. DNAs from isolated clones (0.1–0.3 μg each) were digested with *Eco*RI and electrophoresed in 0.5% agarose gels. The fragments were transferred to nitrocellulose filters[19] and hybridized with ^{32}P-labelled probes as described elsewhere[17]. Ethidium bromide stains (*a*) and autoradiograms (*b*) of Southern blots are shown. In *A*, 4.0-kilobase (kb) *Xba*I–*Hha*I fragment of γ2b gene (Fig. 1*d*) was used as a probe. Lanes 1, 2 and 3 contain DNAs of clones 2, 9 and 32, respectively. In *B*, γ2a cDNA clone (pG2a-14) (probe *e* in Fig. 1) was used as a probe. Lanes 1 and 2 contain DNAs of the 6.6-kilobase γ2b gene clone (IgH22)[28,30] and clone 9, respectively.

To determine directly the order of C$_H$ genes and test the proposed C$_H$ gene organization, we and others cloned mouse immunoglobulin genes, and have determined the complete nucleotide sequences of the γ1, γ2b and μ genes[27–31]. Liu *et al.*[32] isolated a clone which contains both μ and δ genes, and demonstrated that the μ gene is located ~4.5 kilobases 5′ to the δ gene. The linkage of 5′–γ1–γ2b–3′ was shown by cloning overlapping chromosomal segments[20]. We have recently cloned the γ3 and ε genes[33,34], and using these cloned DNA segments as probe, cloned overlapping chromosomal segments from embryonic mouse DNA. Characterization of these clones unequivocally demonstrates that the mouse immunoglobulin γ chain genes are aligned as proposed[12] (that is, 5′–γ1–γ2b–γ2a–ε–3′) on a chromosome. The linkage of 5′–ε–α–3′ is described elsewhere[34].

The γ2b gene is located 5′ to the γ2a gene

When we screened a Charon 4A library containing partial *Hae*III digests of embryonic mouse DNA with the 4-kilobase *Xba*I–*Hha*I fragment (probe *d* in Fig. 1) of the γ2b gene clone[28,30] as a probe, we isolated three clones which were designated Ch · M · Igγ2b–2 (clone 2), Ch · M · Igγ2a–9 (clone 9) and Ch · M · Igγ2a–32 (clone 32). These were characterized by restriction enzyme cleavage mapping and Southern blot hybridization (Figs 1, 2). Clone 2 contains an *Eco*RI fragment (5.6 kilobases) that hybridizes with probe *d* (Fig. 2*A*, lane 1) and shares common restriction sites with the 6.6-kilobase *Eco*RI fragment of the γ2b gene, the nucleotide sequence of which has already been determined[30]. Comparison of restriction maps of clone 2 and the γ2b clone indicates that clone 2 contains the γ2b gene at the 3′ end (Fig. 1).

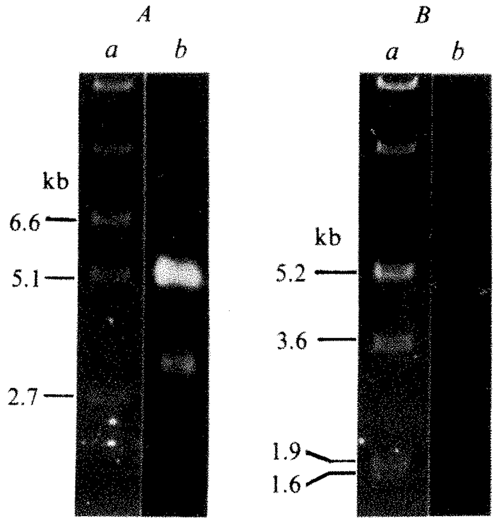


Fig. 3 Cross-hybridization between S$_{\gamma 1}$ and S$_{\gamma 2a}$ regions. Experiments were done as described in Fig. 2 legend. DNA was digested with *Eco*RI and Southern blot hybridized with ^{32}P-labelled 5′ *Eco*RI fragment (5.1 kilobases) of clone 3. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. *A*, clone 3 DNA. *B*, clone 9 DNA. A faint band of 3.1 kilobases in lane *A*(*b*) is due to deletion of the 5.1-kilobase fragment of clone 3 during propagation.

However, restriction cleavage site mapping indicates that clone 9 does not contain any region corresponding to the γ2b structural gene (Fig. 1), but has one *Eco*RI fragment (5.2 kilobase) that hybridizes with the γ2b gene fragment (probe *d*) as shown in Fig. 2*A*, lane 2. Because the γ2b and γ2a genes are known to share extensive homology and cross-hybridize[30,35,36], we thought that clones 9 and 32 could contain the γ2a gene.

To test this we used as probe a γ2a cDNA clone pG2a-14 (probe *e* in Fig. 1) which does not hybridize with the γ2b gene in the stringent conditions used[17]. As shown in Fig. 2*B*, clone 9 was shown to contain an *Eco*RI fragment (5.2 kilobases) that hybridized to the γ2a gene probe (probe *e*). This *Eco*RI fragment is located at the 3′ end of the clone 9 insert and is cleaved by *Xho*I (data not shown). We have determined a partial nucleotide sequence of the 5.2-kilobase *Eco*RI fragment of clone 9. The nucleotide sequence from the *Bam*HI site towards the *Xho*I site of the γ2a coding region is

GGATCCCTGTCCAGTGGTGTG
CACACCTTCCCAGCTGTCCTGCAG,

which coincides with that reported for the CH1 domain of the γ2a cDNA clone[36]. The amino acid sequence

Gly-Ser-Leu-Ser-Ser-Gly-Val-His-
Thr-Phe-Pro-Ala-Val-Leu-Gln

predicted from the nucleotide sequence agrees well with that previously determined for the CH1 domain of the γ2a chain of MOPC173 (ref. 37). The sequence determines the direction of transcription of the γ2a gene and unequivocally demonstrates that clone 9 contains the γ2a gene at the 3′ side.

The restriction map analysis indicates that clones 9 and 32 are almost identical to each other except that clone 32 extends 1.7 kilobases further in the 5′ direction and has a small deletion (0.5 kilobases) in the middle (Fig. 1). It is clear that clone 32 has a γ2a gene fragment at the 3′ end identical to that of clone 9. In addition, clone 32 contains, at the 5′ end, a 3.6-kilobase *Eco*RI fragment which hybridizes with the γ2b gene probe (Fig. 2*A*, lane 3). The restriction map of the 5′-terminal portion (~2.6 kilobases) of clone 32 is indistinguishable from that of the 3′ end of clone 2 (Fig. 1). These results indicate that clone 32 shares overlapped regions at the 5′ end with clone 2 and at the 3′ end with clone 9. The linear alignment of clones 2, 32 and 9 clearly demonstrates that the γ2b gene is located 5′ to the γ2a gene and the two genes are 15 kilobases apart (Fig. 1).

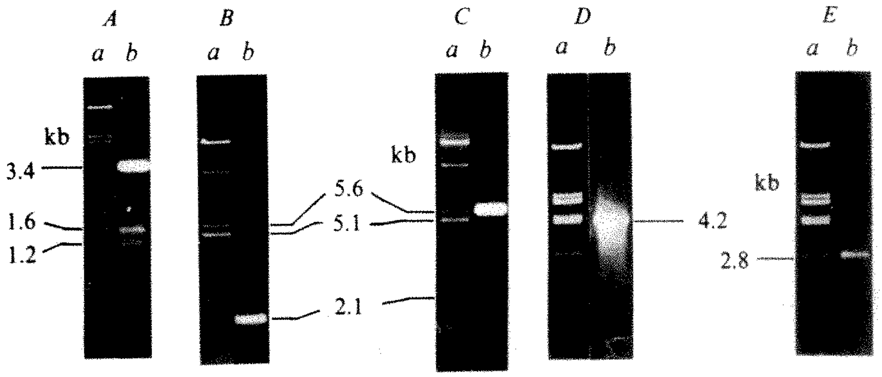


Fig. 4 Southern hybridization of isolated clones linking between the $\gamma1$ and $\gamma2b$ genes and between the $\gamma2a$ and ε genes. Experiments were done as described in Fig. 2 legend. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. In *A*, clone 3 DNA was digested with *Eco*RI and *Bam*HI, and the 6.6-kilobase *Eco*RI fragment of the $\gamma1$ gene (Fig. 1*a*) was used as a probe. In *B*, clone 69 DNA was digested with *Eco*RI and 4.0-kilobase *XbaI–HhaI* fragment of the $\gamma2b$ gene (Fig. 1*d*) was used as a probe. In *C*, clone 69 DNA was digested with *Eco*RI and 2.7-kilobase 3′ *Eco*RI fragment of clone 3 (Fig. 1*b*) was used as a probe. In *D*, clone 12 DNA was digested with *Eco*RI and *Bam*HI, and the probe was 2.2-kilobase *Bam*HI–*Hin*dIII fragment of the ε gene (Fig. 1*h*). *E*, clone 12 DNA was digested with *Eco*RI and *Bam*HI and the probe was the 5.2-kilobase *Eco*RI fragment of clone 9 (Fig. 1*f*).

The $\gamma1$ gene is located 5′ to the $\gamma2b$ gene

The embryonic mouse DNA library was screened with a 6.6-kilobase $\gamma1$ gene fragment (IgH2) cloned previously[27] as a probe (probe *a* in Fig. 1) and two different clones were obtained: Ch · M · Igγ1 – 3 (clone 3) and Ch · M · Igγ1 – 17 (clone 17). Restriction maps of these are shown in Fig. 1.

Comparison of the restriction maps of clones 3 and 17, and the $\gamma1$ gene clone (IgH2) indicates that both clones 3 and 17 contain the 6.6-kilobase *Eco*RI fragment of the $\gamma1$ gene. The comparison also allows determination of the orientation of the two clones. The insert of clone 17 is $\sim$2 kilobases shorter than that of clone 3 and the deletion is located in the 5′-terminal *Eco*RI fragment. These two clones may be derived from a common ancestral phage and the deletion may have occurred during propagation of phages, because the directions of the inserts within these phages and the ends of the inserts in these clones are identical to each other. Furthermore, clone 3 underwent deletion during propagation, producing a phage similar to clone 17 (see Fig. 3*A*). Southern blot experiments (Fig. 4*A*) provide additional evidence that clone 3 contains the $\gamma1$ gene. When clone 3 DNA was digested with *Eco*RI and *Bam*HI simultaneously, clone 3 DNA produced 3.4-, 1.6- and 1.2-kilobase fragments which hybridize to the $\gamma1$ gene (probe *a*), as expected from the restriction map.

Because $\gamma1$ gene clones (3 and 17) and $\gamma2b$ gene clone (2) do not share a common region, we screened the embryonic mouse DNA library with a 3′ fragment of clone 3 (Fig. 1, probe *b*) and a 5′ fragment of clone 2 (Fig. 1, probe *c*) as probes. We isolated a new clone designated Ch · M · Igγ2b–69 (clone 69)—the restriction map of this is shown in Fig. 1.

The 3′ part of clone 69 is indistinguishable from that of clone 2 except that the former has an $\sim$1.5-kilobase deletion at the 5′ side of the $\gamma2b$ gene. Comparison of restriction maps of clones 69 and 2 indicates that clone 69 contains the $\gamma2b$ gene at the 3′ end. Southern blot hybridization experiments show that a 2.1-kilobase *Eco*RI fragment of clone 69 hybridizes with the $\gamma2b$ gene probe (Fig. 4*B*). Clone 69 also contains a 5.6-kilobase *Eco*RI fragment that hybridizes with the 3′-terminal fragment of clone 3 (Fig. 1, probe *b*) as shown in Fig. 4*C*. From these results it is clear that the 3′ region (2.1 kilobases) of clone 3 and the 5′ region (2.1 kilobases) of clone 69 are identical. The linear alignment of clones 3, 17, 69 and 2 clearly demonstrates that the $\gamma1$ gene is located 5′ to the $\gamma2b$ gene and the two genes are 21 kilobases apart. These results are consistent with those of Maki *et al.*[20].

The ε gene is located 3′ to the $\gamma2a$ gene

We recently cloned a $\gamma2a$ gene fragment from *Eco*RI digests of RPC5 myeloma DNA designated Ch · M · Igγ2a–11 (clone 11). Southern blot hybridization and R-loop mapping show that the $\gamma2a$ gene is located at the 5′ end of the 20-kilobase cloned fragment (data not shown). The restriction map of clone 11 is shown in Fig. 1. The 5′ region of the RPC5 $\gamma2a$ gene fragment shares common restriction sites with the 3′ region of clones 32 and 9.

The 8.5-kilobase *Hin*dIII fragment (probe *g*) of clone 11 was recloned into pBR322 and used as a probe to screen the mouse embryonic DNA library. We obtained a germ-line $\gamma2a$ gene fragment designated Ch · M · Igε–12 (clone 12)—the restriction map of this is also shown in Fig. 1. The 16.5-kilobase insert of clone 12 has restriction sites almost identical to those of the 3′ portion of clone 11 except that the latter has a deletion at the 3′ part.

Cloning of an expressed ε gene from an ε chain-producing hybridoma has been reported elsewhere[34]. Partial nucleotide sequence determination indicates that our ε gene clone Ch · M · Igε–1 encodes an amino acid sequence similar to that of human ε chain. The expressed ε gene (Ch · M · Igε–1) contains, as do other expressed H chain genes[20–26], at least three germ-line gene segments: a V gene, a J gene (including S_μ region that flanks the 5′ end of the μ gene) and the ε gene. The restriction map of the region encoding the ε gene of Ch · M · Igε–1, shown at the top of Fig. 1, seems to be identical to that of the 3′ region of clone 12.

To test whether or not clone 12 contains the ε gene, Southern blot hybridization experiments were done using a *Bam*HI–*Hin*dIII fragment (probe *h*) of the ε gene as a probe. Figure 4*D* shows that probe ε hybridizes a 4.2-kilobase *Bam*HI + *Eco*RI fragment at the 3′ part of clone 12. We also confirmed (Fig. 4*E*) that the 5′ segment of clone 12 has a region identical to the 3′ segment (probe *f*) of clone 9 which contains the germ-line $\gamma2a$ gene. Taken together, these results indicate that the ε gene is located 3′ to the $\gamma2a$ gene and the two genes are about 15 kilobases apart.

We have recently isolated three clones[38] which link the $\gamma1$ gene (clone 3) and the $\gamma3$ gene (Igγ3–30)[33]. Our data indicate that the $\gamma3$ gene is located 34 kilobases 5′ to the $\gamma1$ gene, and we

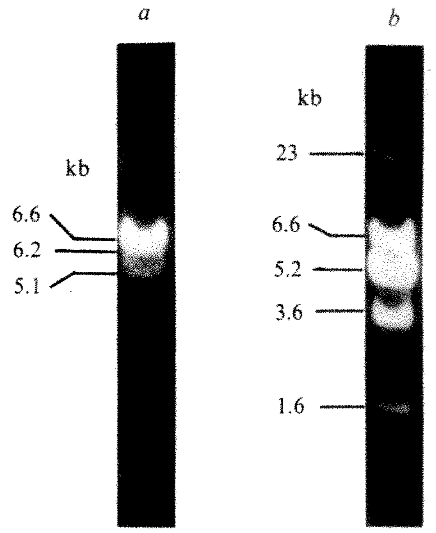


Fig. 5 Distance between the $\gamma1$, $\gamma2b$ and $\gamma2a$ genes. BALB/c newborn mouse DNA (3 μg) was digested with *Eco*RI to completion and run on 0.5% agarose gel electrophoresis. DNA was transferred to a nitrocellulose filter[19] and hybridized with ^{32}P-labelled probes as described elsewhere[17]. Total DNAs of clones 69 and 9, respectively, were used as probes in *a* and *b*.

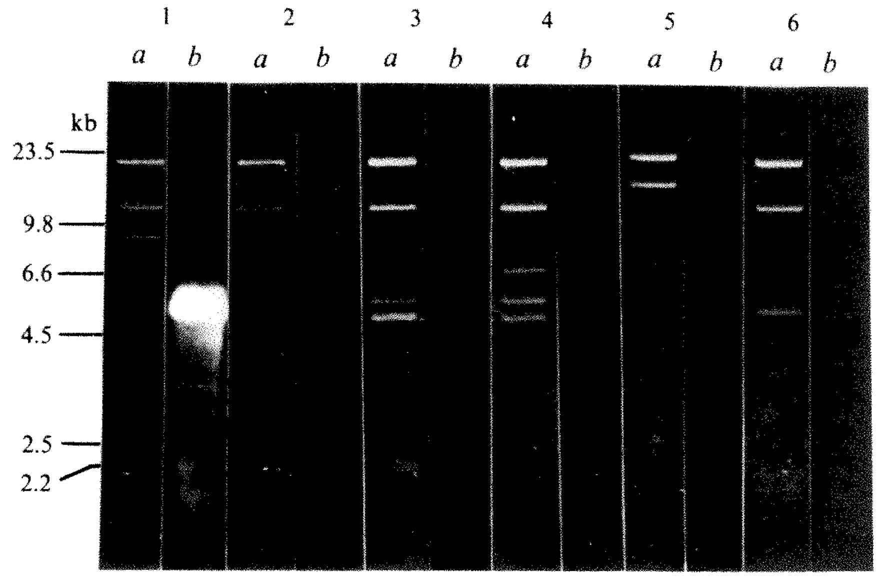


Fig. 6 Hybridization analysis with J-region probe. DNAs from isolated clones covering the region between the $\gamma1$ and $\gamma2a$ genes were digested with *Eco*RI and run on 0.7% agarose gel electrophoresis. DNAs were transferred to a nitrocellulose filter[19] and hybridized with the 1.5-kilobase *Hin*dIII–*Eco*RI fragment that is located 6.5 kilobases 5′ to the μ gene. This fragment contains the J_{H4} region[24]. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. DNA in each lane is as follows: (1) the expressed $\gamma1$ gene of MC101 myeloma (Igγ1–706)[25,47,48] that contains V, D, J_3, J_4, S_μ and $C_{\gamma1}$ gene segments; (2) clone 3; (3) clone 69; (4) clone 2; (5) the 6.6-kilobase $\gamma2b$ gene clone (IgH22)[28,30]; (6) clone 9.

have demonstrated[34] that the ε gene is present 12.5 kilobases upstream of the α gene. The linkage of J_H–μ–δ has been shown elsewhere[24,32]. Thus it is now clear that the organization of the mouse C_H gene is 5′–J_H–(6.5 kilobases)–μ–(4.5 kilobases)–δ–(unknown distance)–$\gamma3$–(34 kilobases)–$\gamma1$–(21 kilobases)–$\gamma2b$–(15 kilobases)–$\gamma2a$–(14.5 kilobases)–ε–(12.5 kilobases)–α–3′. Note that all the C_H genes ordered have the same orientation and thus are transcribed from the same strand of DNA.

Distance between C_H genes in germ-line DNA

Comparison of the restriction maps of the clones isolated indicate that some have deletions which probably occurred during propagation of phages. This was also reported for globin gene clones[39]. We therefore compared the sizes of *Eco*RI fragments obtained from cloned γ genes and from Southern blot hybridization of newborn mouse DNA. When the total insert of clone 69 was used as a probe, *Eco*RI digests of newborn mouse DNA produced 6.6(doublet)-, 6.2- and 5.1-kilobase bands which coincide with the *Eco*RI fragments deduced from the cloned DNAs (Fig. 5). When the total insert of clone 9 was used as a probe, *Eco*RI digests of newborn mouse DNA produced 23-, 6.6-, 5.2-, 3.6- and 1.6-kilobase bands which coincide with the *Eco*RI fragments deduced from the cloned DNAs.

Furthermore, consideration of the maximum size of the insert (21 kilobases) that can be taken by Charon 4A phage makes it unlikely that there is a deletion >2 kilobases in the clones obtained. Because the insert sizes of clones 69 and 32 are 19.3 and 19.2 kilobases, respectively, the original inserts cannot be more than 2 kilobases greater than those of the present clones. We therefore conclude that the distances between the $\gamma1$ and $\gamma2b$ genes and between the $\gamma2b$ and $\gamma2a$ genes in the germ-line chromosome are 21.2 and 15.5 kilobases, respectively. Similar experiments using the ε gene as a probe[34] indicate that the distance between the $\gamma2a$ and ε gene is 14.5 kilobases in newborn mouse DNA.

Absence of J region sequences in the entire region encompassing γ and ε genes

We determined whether or not each γ gene clone contains a sequence homologous to a J region. We used a J segment (J_4) as a probe and hybridized it with restricted DNA fragments of all the γ gene clones shown in Fig. 1. There were no J or J-like sequences between 8 kilobases 5′ to the $\gamma1$ gene and 3 kilobases 3′ to the $\gamma2a$ gene (Fig. 6). The J_4 sequence was actually expressed in the $\gamma2b$ chain produced by MOPC141 myeloma[24]. It can be argued that other completely different J sequences may be present in the γ gene-flanking regions. If this is the case, V region sequences formed by the use of J_μ and J_γ are different. This is unlikely because the same idiotype (a V region sequence) can be expressed as either μ or γ chain. It is possible that a γ gene has completely different J-flanking regions and an identical J_4 sequence of 51 base pairs which was not detected by our probe. This also seems unlikely because the precise joining between the same V gene and J_γ probably requires a nucleotide sequence similar to that of the J_μ-flanking region[23,24]. A similar experiment was done on the ε gene clone (clone 12) and a $\gamma3$ gene clone[33], in which we did not find a J region or J-like sequences (data not shown).

The S region before each C_H gene may comprise repetitive sequences

The class-switch recombination takes place at the region 5′ to each C_H gene. Such regions responsible for the class switch have been functionally defined as S regions[21]. We have recently demonstrated that the $S_{\gamma1}$, $S_{\gamma2b}$ and $S_{\gamma3}$ regions comprise tandem repetitive conserved sequences of 49 base pairs[25,33]. The 5′-flanking region of the $\gamma2b$ gene which contains the repetitive sequence is deleted from clone 69. Such tandem repetitive sequences of conserved 49-base pair units can be easily deleted when propagated in bacteria[40]. Deletions of short segments (0.5–2.0 kilobases) found in several clones are always located in the 5′-flanking region of each γ gene. The distance from the 5′ end of each γ gene to a deleted region is almost constant (~3 kilobases). During propagation of clone 9, we also found a deletion of ~1.5 kilobases at ~3 kilobases 5′ of the $\gamma2a$ gene, about the same place where deletion occurred in clone 32. The deletions in these clones may indicate the positions of repetitive sequences.

Although we have not determined the nucleotide sequence of the $S_{\gamma2a}$ region, the deletion in clone 32 may suggest a similar repetitive sequence in the 5′-flanking region of the $\gamma2a$ gene. We found that the 5′-flanking region of the $\gamma2a$ gene cross-hybridizes with the $S_{\gamma1}$ region (Fig. 3). The S_μ and S_α regions also contain repetitive sequences[22,31,41] and these are often deleted during cloning. Thus it is likely that each C_H gene, except for the δ gene which is probably expressed by RNA splicing[32], has its own S region for class-switch recombination and that all the S regions comprise repetitive sequences.

We thank Dr P. Leder for an embryonic mouse DNA library, Dr K. Matsubara for critical reading of the manuscript, Mrs S. Nishida for technical assistance, and Miss T. Wada for preparation of the manuscript. Research was supported in part by grants from the Ministry of Education, Science and Culture of Japan and from Asahi Scientific Research Fund.

Received 30 September; accepted 24 November 1980.

1. Landucci-Tosi, S., Mage, R. G. & Dubiski, S. *J. Immun.* **104,** 641–647 (1970).
2. Kindt, T. J., Mandy, W. J. & Todd, C. W. *Biochemistry* **9,** 2028–2032 (1970).
3. Knight, K. L. & Hanly, W. C. *Contemp. Topics molec. Immun.* **4,** 55–88 (1975).
4. Nossal, G. J. V., Warner, N. L. & Lewis, H. *Cell Immun.* **2,** 41–53 (1971).
5. Fu, S. M., Winchester, R. J. & Kunkel, H. G. *J. Immun.* **114,** 250–252 (1975).

6. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1707–1711 (1975).
7. Goding, J. W. & Layton, J. E. *J. exp. Med.* **144,** 852–857 (1976).
8. Sledge, C., Fair, D. S., Black, B., Krueger, R. G. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **73,** 923–927 (1976).
9. Bleux, C., Ventura, M. & Liacopoulos, P. *Nature* **267,** 709–711 (1977).
10. Wang, A.-C., Wang, I. Y. & Fudenberg, H. H. *J. biol. Chem.* **252,** 7192–7199 (1977).
11. Kincade, P. W., Lawton, A. R., Bockman, D. E. & Cooper, M. D. *Proc. natn. Acad. Sci. U.S.A.* **67,** 1918–1925 (1970).
12. Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2140–2144 (1978).
13. Rabbitts, T. H., Forster, A., Dunnick, W. & Bentley, D. L. *Nature* **283,** 351–356 (1980).
14. Coleclough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1422–1426 (1980).
15. Cory, S. & Adams, J. M. *Cell* **19,** 37–51 (1980.)
16. Cory, S., Jackson, J. & Adams, M. *Nature* **285,** 450–456 (1980).
17. Yaoita, Y. & Honjo, T. *Biomed. Res.* **1,** 164–175 (1980).
18. Yaoita, Y. & Honjo, T. *Nature* **286,** 850–853 (1980).
19. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
20. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2138–2142 (1980).
21. Kataoka, T., Kawakami, T., Takahashi, N. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **77,** 919–923 (1980).
22. Takahashi, N., Kataoka, T. & Honjo, T. *Gene* **11,** 117–127 (1980).
23. Early, P. W., Huang, H., Davis, M. M., Calame, K. & Hood, L. *Cell* **19,** 981–992 (1980).
24. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
25. Honjo, T. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
26. Davis, M. H. *et al. Nature* **283,** 733–739 (1980).
27. Honjo, T. *et al. Cell* **18,** 559–568 (1979).
28. Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4240–4244 (1979).
29. Tucker, P. W., Marcu, K. B., Newell, N., Richard, J. & Blattner, F. R. *Science* **206,** 1301–1306 (1979).
30. Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. & Honjo, T. *Nature* **283,** 786–789 (1980).
31. Kawakami, T., Takahashi, N. & Honjo, T. *Nucleic Acids Res.* **8,** 3933–3945 (1980).
32. Liu, C.-P., Tucker, P. W., Mushinski, J. F. & Blattner, F. R. *Science* **209,** 1348–1353 (1980).
33. Kataoka, T., Miyata, T. & Honjo, T. *Cell* (in the press).
34. Nishida, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
35. Yamawaki-Kataoka, Y. *et al. Biochemistry* **18,** 490–494 (1979).
36. Sikorav, J.-L., Auffray, C. & Rougeon, F. *Nucleic Acids Res.* **8,** 3143–3155 (1980).
37. Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J. & Schiff, C. *Annls Immun. Inst. Pasteur, Paris* **127,** 607–631 (1976).
38. Takahashi, N., Shimizu, A. & Honjo, T. (in preparation).
39. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19,** 959–972 (1980).
40. Tsujimoto, Y. & Suzuki, Y. *Cell* **16,** 425–436 (1979).
41. Davis, M. M., Kim, S. K. & Hood, L. E. *Science* **209,** 1360–1365 (1980).
42. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
43. Blattner, F. R. *et al. Science* **196,** 161–169 (1977).
44. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
45. Maniatis, T., Jeffrey, A. & Kleid, D. K. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
46. Tiemeier, D. C., Tilghman, S. M. & Leder, P. *Gene* **2,** 173–191 (1977).
47. Obata, M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
48. Kataoka, T. *et al.* (in preparation).

LETTERS

Variability of the double quasar 0957+561 A, B

Joseph S. Miller, Robert R. J. Antonucci & William C. Keel

Lick Observatory, Board of Studies in Astronomy and Astrophysics,
University of California, Santa Cruz 95064

Evidence suggests that the double nature of the quasar 0957+561 A, B results from the influence of an intervening gravitational lens on the light from a more distant, single quasar. The nearly identical spectral characteristics of the two objects have been established[1-3] and have been the incentive for interpreting the two apparent objects as images of a single one. We report here spectroscopic and photometric observations which show that the images of the double quasar 0957+561 A, B are variable. Brightness changes were observed to be accompanied by corresponding spectral variations. These data suggest that the quasar itself is an intrinsically variable object.

Photographs by one of us (W.C.K.) taken with the 0.9-m Lick Observatory Crossley reflector on 18 and 19 May 1980 (all dates are given in UT) showed that ratio of brightness in the blue of B to A was 0.95, considerably different from the value of about 0.75 given earlier[1,4] (A is taken as northern image) and it seemed more likely that B brightened than that A faded[5]. To investigate further the nature of the variability, we obtained additional Crossley photographs and made spectrophotometric observations with the Lick Observatory 3-m Cassegrain image-tube scanner on 9 June 1980. We also obtained spectrophotometric data for the pair on 12, 13 and 14 March 1980 using a similar observing set-up. The instrument and observing procedures have been described elsewhere[6].

Plates were obtained with the Crossley reflector on 19 May and 7 and 21 June 1980. The plate-filter combination, 103a-0 plus GG 13, corresponds approximately to the B bandpass. We also have a plate taken on 20 March with the Lick Observatory 20-inch astrograph for the proper motion program which recorded the double quasar by chance; this plate had roughly the same spectral response as the Crossley plates. Lacking the extension of a photoelectrically calibrated sequence of stars to magnitudes comparable with those of the quasar images, we present here only flux ratios. The ratios, given in Table 1, were determined from IRIS photometry of the plates. We extrapolated the slope of the IRIS photometry calibration to the brightness of the double quasar, which was ~2 mag fainter than our faintest calibrated star. Only a small additional error is introduced by this procedure when measuring ratios close to unity, and this is included in the errors quoted in Table 1. These ratios refer almost entirely to the continuum fluxes, because the only strong line included, C III] λ 1,909, has a small equivalent width compared with the total bandpass of the plate-filter combination. Also, these measures are only very slightly affected by the presence of the lens galaxy near component B. All measures used comparable IRIS openings corresponding to ~3 arc s diameter, thus including much of the galaxy light with the southern quasar image. This introduced a constant additional contribution, which, from the red magnitude given by Young *et al.*[4] and an approximate *K*-correction, we estimate to be ~3% of the image B flux; this small correction is not

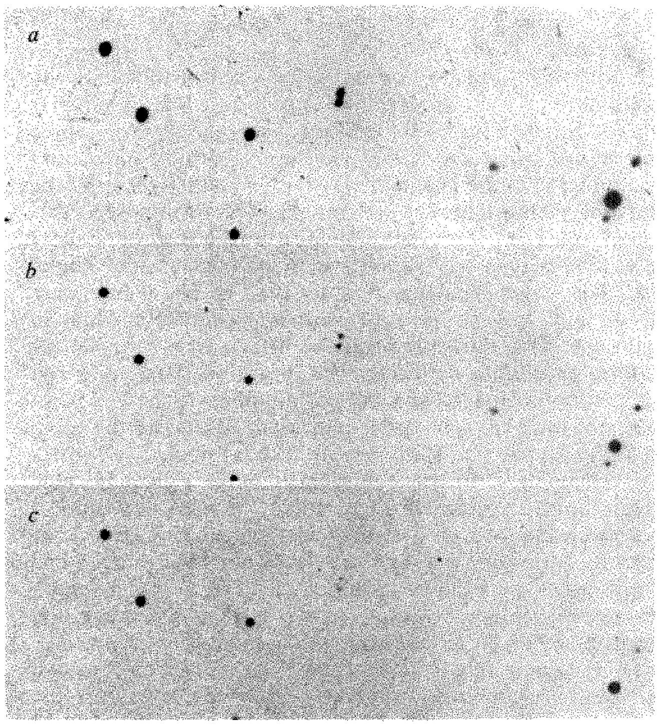


Fig. 1 Three photographs of the double quasar taken with the Lick Observatory Crossley reflector. *a*, 22 January 1915: the images are slightly comatic because the double was not centred on the original plate. *b*, 18 May 1980. *c*, 21 June 1980: this plate was underexposed because of poorer observing conditions, but shows B brighter than A. In all pictures north is at the top and east is at the left, so that image A is the upper of the double; the double is close to the centre of each of the three.

0028-0836/81/030153—03$01.00

Table 1 0957+561 A, B brightness ratios (B/A)

Date (UT)	Ratio
20 March 1980	0.84±.07
19 May 1980	0.95±.05
7 June 1980	1.15±.07
21 June 1980	1.09±.07

important for the results in Table 1 and has not been included. All exposures were well below the level at which saturation and growth of the galaxy image could affect the measured magnitudes. Figure 1 reproduces two of the 1980 Crossley plates. Observing conditions were less satisfactory on 7 June, resulting in a plate which reproduced poorly, but the images contain sufficient information to provide a reliable flux ratio.

Lick Observatory files contain Crossley plates taken in 1902, 1903 and 1915 which recorded the double quasar, with the plates centred on the galaxy NGC3079. Figure 1 reproduces the plate taken on 22 January 1915 which shows the brightness ratio of the double near that given previously[1,3,4]. The other archival Crossley plates exhibit a similar flux ratio for the pair, but preliminary analysis indicates that both images were systematically brighter in the past. A detailed analysis of the past history of the images on the archival plates is in progress and will be presented elsewhere.

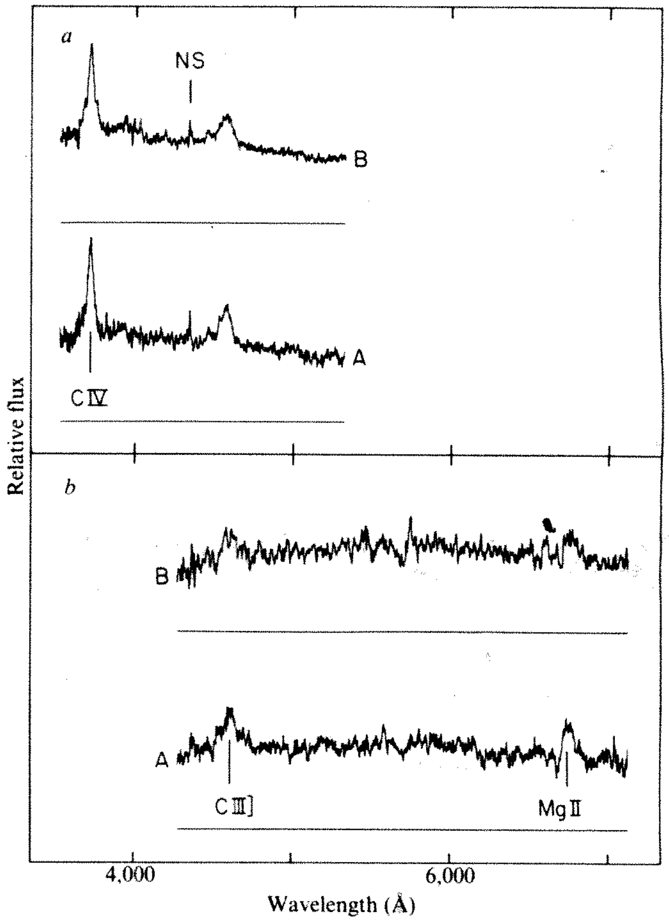


Fig. 2 Spectra of the double quasar. *a*, 12–14 March 1980; *b*, 9 June 1980. The total observing time for each of the March spectra was ~2.5 h, but observing time limitations allowed only 32 min for each in June, resulting in noisier spectra from the June data. A and B had very similar spectra in March, but in the June spectra the equivalent widths of the emission lines are smaller for B than for A. NS marks the position of a partially uncancelled strong night sky emission line.

Additional information on the variability is available in the spectrophotometry. Because we used a 2.8 arc s circular entrance aperture for both the March and June spectrographic observations to reduce the sky contribution and avoid contamination of the image being observed with light from the other image, the data are not reliable for absolute spectrophotometry. The March data show that image B was definitely fainter than image A and are consistent with the result of Young *et al.*[4] that A was ~0.3 mag brighter than B. The 9 June data indicate that the two images were approximately equal in brightness. The data obtained in March are suitable for measurements of the equivalent widths of the C IV λ1,549 and C III] λ1,909 emission lines. We did not observe the UV spectrum on 9 June but can derive equivalent widths for the C III] and Mg II λ2,800 emission lines from those data. The spectra are shown in Fig. 2, and the equivalent widths are given in Table 2. To produce an internally consistent set of equivalent widths for a given line, we used an identical wavelength interval to define the extent of the line profile on all of the individual spectra. Each spectrum is actually derived from two independent spectra. The spectrograph has two entrance apertures, and an object under observation is switched between them, ultimately spending equal time in both. The data from the two apertures are reduced separately. Comparisons of the equivalent widths determined from the independent spectra and similar data for other objects indicate that the errors in the equivalent widths are unlikely to exceed 5Å. Although we include values from earlier workers, we have not compared their values with ours because of the possibility of systematic differences between groups in methods of evaluating equivalent widths in data of only moderate signal-to-noise ratio. In our March data, the equivalent widths of C IV and C III] in images A and B agreed to within the measuring errors, as the earlier workers found. However, the June data show a marked change, as C III] and Mg II have significantly smaller equivalent widths in B than they do in A. Also, for both objects, the June equivalent widths of C III] differ from those measured from the March data, though the change for image A is only of marginal significance given the estimated errors. A simple interpretation of these data is that between March and June, the continuum of A faded slightly, the continuum of B brightened substantially, and all emission lines remained constant in flux. This spectroscopically motivated conclusion is in complete accord with the photographic photometry mentioned above. Because no change in the properties of the gravitational lens would be expected to produce changes in the equivalent widths of various lines, we conclude that 0957+561 is an intrinsically variable quasar, and the spectral differences between the two images in June result from different light-travel times for the rays responsible for each image. Other studies of variable QSOs (see for example, refs 7, 8) indicate that, at least on short time scales (≤1–10 yr), only continua have been observed to vary while line fluxes remained

Table 2 Equivalent widths (Å) in the observed frame

	0957+561 A			0957+561 B		
	C IV λ1,549	C III] λ1,909	Mg II λ2,800	C IV λ1,549	C III] λ1,909	Mg II λ2,800
12, 13 March	60	47	—	55	48	—
9 June	—	58	41	—	32	28
WCW[1]	68	54	28:	70	55	Present
WW[3]	—	58	61	—	58	62
Young *et al.*[4]	77	62	58	74	58	45

constant. Furthermore, the ratio of equivalent widths derived from the June data of C III] for the two images is different in size from that of Mg II, but this difference is consistent with a considerably greater galaxy contribution to the continuum of B in the spectral region of Mg II than in the region of C III], as was pointed out by Young *et al.*[4]

The double quasar 0957+561 A, B is, therefore, variable in both brightness and spectrum. The light variations were accompanied by corresponding changes in the equivalent widths of emission lines, and we conclude that the quasar is an intrinsic variable with behaviour similar to that observed in other variable quasars. Establishing a value for a constant time-difference between variations of the two images would not only confirm the gravitational lens hypothesis, but would add an additional parameter to models for the lens. As Young *et al.*[4] have pointed out, it is not even clear which image is formed by the path of shortest light-travel time.

We thank G. D. Schmidt for assistance in the observations. This research was supported in part by NSF grant AST 76-20843. This is Lick Observatory Bulletin no. 824.

Received 15 August; accepted 11 November 1980.

1. Walsh, D., Carwell, R. F. & Weymann, R. J. *Nature* **279,** 381–384 (1979).
2. Weymann, R. J. *et al. Astrophys. J. Lett.* **233,** L43 (1979).
3. Wills, B. J. & Wills, D. *Astrophys. J.* **238,** 1–9 (1980).
4. Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. S. *Astrophys. J.* (in the press).
5. *IAU Circ.* No. 3481 (1980).
6. Miller, J. S., Robinson, L. B. & Wampler, E. J. in *Advances in Electronics and Electron Physics*, 693–698 (Academic, New York, 1976).
7. Netzer, J., Wills, B. J. Uomoto, A. E., Rybiski, P. M. & Tull, R. G. *Astrophys. J. Lett.* **232,** L155–L159 (1979).
8. Miller, J. S. & French, H. B. in *Proc. Pittsburgh Conf. on BL Lac Objects*, 228–232 (University of Pittsburgh, 1978).

An outburst of comet Tempel-2 observed spectrophotometrically

Paul E. Johnson*, Dale W. Smith† & Richard W. Shorthill‡

* NASA-NRC Resident Research Associate, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103
† Astronomy Department, University of Washington, Seattle, Washington 98195
‡ Geospace Sciences Laboratory, University of Utah Research Institute, Salt Lake City, Utah 84108

Periodic comet Tempel-2, first discovered in 1873, is a possible target of a spacecraft flyby in the 1980s. Accurate physical measurements of this comet have only recently been attempted. Few spectrophotometric data exist for any faint comet, but we present here such data for Tempel-2 at a heliocentric distance of 2.97 AU. The data were obtained during an outburst in which the comet exhibited heightened CN and C_3 emission over a quiescent phase. Comet flares of such magnitude are exceptional, and are particularly surprising for a periodic comet at such a large heliocentric distance.

We observed Tempel-2 on 22.2 December 1978 UT using the Oke multichannel spectrometer[1] and 200-inch telescope. Fluxes of the comet and the sky 40 arc-s away were measured simultaneously in 30 channels in the wavelength range 3,200–10,500 Å. The bandwidth was 160Å below 5,720 Å and 360 Å above (in rectangular bandpasses). All measurements were made using a 14-arc s aperture. The air mass of the comet was 1.1 during the observations.

Tempel-2 was post-perihelion at a heliocentric distance of 2.97 AU with a phase angle of 8.61° and a magnitude at 5,600 Å of ~18.5. Identification at the telescope was confirmed by motion of the comet on the television monitor with the expected rate and direction. No coma was evident on the monitor, but the seeing was poor.

The sky background was removed by differencing the count rates in the spectrometer apertures on the comet and the sky. The relative sensitivity of the apertures was determined by observing the same object in each aperture. The spectral reflectance was obtained by taking the ratio, in each channel, of the flux due to the comet and that of HD28992, a G1V star in the Hyades observed at nearly the same air mass. (A Hyades star was chosen because it would be expected to have a nearly solar metal abundance and thus provide the closest possible spectral match to the Sun.)

The accuracy of this method was confirmed by applying it to measurements of the bright asteroid 14 Irene. An excellent match with the data of Chapman and Gaffey[2] was obtained within the confidence limits of their data (±3%). The spectral reflectance data for the comet are listed in Table 1 and plotted in Fig. 1, including channels from 3,480 Å to 8,160 Å. Channels at longer and shorter wavelengths are very noisy and are not shown.

Comparison of our data with the photometry of other observers (ref. 3 and R. L. Millis, personal communication) shows that our observations were made during an outburst in which the continuum brightness of this comet temporarily increased by about one visual magnitude. This suggests that the reflectance spectrum is produced principally by scattering of sunlight off particles around the nucleus and is not diagnostic of the nucleus itself. An additional outburst of ~1.5 mag was observed by Zellner *et al.*[4] a month later with broadband photometry. Although one flare[5] of Tempel-2 was reported during its 1946 apparition, to our knowledge these are the first confirmed outbursts of Tempel-2.

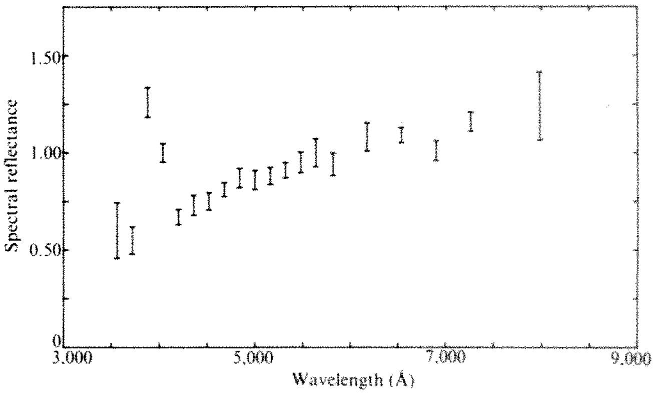


Fig. 1 The spectral reflectance of Tempel-2, normalized to unity at 5,640 Å. These measurements were made on 22.2 December 1978 UT when the comet was post-perihelion at a heliocentric distance of 2.97 AU. The data are indicated with ± 1σ confidence intervals.

The occurrence of a series of outbursts at such a large heliocentric distance is surprising. Only one other short-period comet, Schwassman–Wachmann 1 (the heliocentric distance of which varies between 5.5 and 7.3 AU), has been reported to exhibit high-amplitude outbursts at heliocentric distances ≳3 AU. Note that Schwassman–Wachmann 1 is not a typical short-period comet for several reasons[6]. First, this comet undergoes frequent non-periodic variations in its brightness by a factor of 100. Second, there has been no confirmed measurement of molecular emission from Schwassman–Wachmann 1 other than a weak CO^+ feature[7,8]. Observations of a 1–2 mag flare in any comet are quite rare, even near perihelion where such effects would be expected to be observed most readily, due to both the increased brightness and expected increase in activity of comets at perihelion. Richter[9] has looked for anomalous outbursts of every comet observed photometrically between 1880 and 1947. Of these 358 comets, only 12 had observed flare-ups in total light of a magnitude comparable with that which we measured for Tempel-2 (~1 mag).

---

* Present address: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK.
† Present address: Department of Physics and Astronomy, Western Washington University, Bellingham, Washington 98225.

With the measured spectral reflectance of Tempel-2, we can estimate the abundance of gas and dust in the coma and the albedo of the dust. The emission bands in the channels at 3,880 Å and 4,040 Å are due to the $\Delta\nu = 0$ band of CN at 3,883 Å and the 4,050 Å C_3 group and permit a determination of the CN and C_3 production rates. The CN flux is taken to be the excess flux, in the channel centred at 3,880 Å, above the continuum represented by a linear interpolation between the spectral reflectances in the adjacent channels at 3,720 Å and 4,200 Å. Using the flux (J. B. Oke, personal communication) of the standard star HD84937, we find that this excess represents a flux in the CN band of 8.8×10^{-18} W m^{-2}. This flux is next converted to a number abundance of CN molecules using the relation between the flux per unit column-number density and the heliocentric Doppler velocity as reported by Tatum and Gillespie[10]. For the geocentric distance of 2.062 AU and an $\dot{r}$ of 9.8 km s^{-1}, the CN abundance in the field of view is 1.6×10^{27} molecules.

In the Haser model[11,12], observed molecular column densities in the steady state case can be related to production rates by

$$D(R) = C[\exp(-R/l_d) - \exp(-R/l_p)]/R^2 \qquad (1)$$

where $D(R)$ is the number density of a given molecule at a distance R from the nucleus, l_p is the scale length for the decomposition of the parent molecule, l_d is the scale length for the destruction of the observed species, and C is a constant of proportionality. Using the scale lengths given for CN by Delsemme[13] and the CN abundance derived above, we find a production rate of 5×10^{23} CN molecules s^{-1}. It is ~100 times greater than the upper limit of 6×10^{21} CN molecules s^{-1} found by Spinrad *et al.*[3] during a post-perihelion quiescent phase two months earlier at a heliocentric distance of 2.66 AU. Similarly we find a C_3 production rate[12] of ~10^{24} C_3 molecules s^{-1}. Note that although the region of sampled sky was less than l_d from the comet nucleus, the R^2 term in equation (1) limits the amount of emission to be expected in the sky channel to ≤10% of the emission in the object aperture, if the Haser model is applicable.

Table 1 Spectral reflectance of Tempel-2

Wavelength (Å)	Bandwidth (Å)	Normalized reflectance
3,560	160	0.60
3,720	160	0.55
3,880	160	1.26
4,040	160	1.00
4,200	160	0.67
4,360	160	0.73
4,520	160	0.75
4,680	160	0.81
4,840	160	0.87
5,000	160	0.86
5,160	160	0.88
5,320	160	0.91
5,480	160	0.95
5,640	160	1.00
5,820	360	0.94
6,180	360	1.08
6,540	360	1.09
6,900	360	1.01
7,260	360	1.16
7,620	360	*
7,980	360	1.24

* Unreliable channel.

Next we investigate the role of dust in contributing to the brightness of the comet through reflection. Following Spinrad *et al.*[3] we assume a CN/H_2O number ratio of 1 : 200 and a dust/gas mass ratio near unity. We then have a dust production rate of 3,000 g s^{-1}. Delsemme and Miller[14] derive a relationship between the total gas vaporization rate of a comet and its dust expansion velocity. Arbitrarily assuming dust expansion in a 60° cone, one finds a dust expansion velocity ~10 m s^{-1} (ref. 3). With a dust particle size distribution characterized by a geometric cross-section of 3.17×10^{-7} km^2 g^{-1} (ref. 15) and the dust production rate given above, we derive a total cross-section of 1.6×10^3 km^2 for the dust situated within the projected boundary of our aperture.

Finally, we can estimate the albedo of the dust by assuming that all the continuum flux is produced by scattering of sunlight off of the dust. Using the cross-section obtained above and a nominal phase coefficient of 0.03 mag deg^{-1} (ref. 3), a geometric albedo of ~6×10^{-4} would be required at 5,600 Å. This contrasts with the geometric albedo of 0.10–0.15 suggested by Spinrad *et al.*[3] for the nucleus. The albedo we derive is actually a lower limit. Assuming a 'burst' model, where all of the molecules and dust contributing to the comet brightness are located inside the projected object aperture, the CN abundance, as before, is 1.6×10^{27} molecules. Again assuming a CN/H_2O number ratio of 1 : 200 and a dust/gas mass ratio near unity, we find 9.2×10^6 g, or 2.9 km^2, of dust within the field of view. This leads us to a geometric albedo ~0.3, in closer agreement to the value obtained by Spinrad *et al.*, and nearer to what one would expect for ice grains.

The burst model requires that the time, t, elapsed since the outburst of matter from the nucleus is less than r/v, where r is the maximum radial distance from the nucleus still within the aperture and v is the thermal velocity of the parent molecule of CN. With a v of 0.71 km s^{-1} (ref. 16) and an r of 1.05×10^4 km, we find that $t \leq 10^4$ s. Since the albedo of 0.3 derived from this model certainly seems to agree within a factor of 10 with what one would expect, these results suggest that this observed outburst of Tempel-2 was of relatively short duration, ≤ 1 day. Such rapid outbursts have in fact been observed in Schwassman–Wachmann 1 and seem to occur in the other comets which have exhibited large brightness fluctuations[9].

The present results suggest rapid variability of Tempel-2. More quantitative results on faint comets are hampered by the extreme paucity of data. In particular, more spectrophotometry of faint comets is needed to constrain comet modelling.

We thank Dr R. L. Newburn Jr for supporting this research (via JPL contract BT-708801) and Professor B. Zellner, Professor E. Barker and Dr J. Degewij for helpful discussions. We thank the staff of Hale Observatories for observing time and help with measurements. This work was supported by NASA under contracts NAS7-100 and NGR 49-002-142. P.J. acknowledges support from the National Research Council during part of this study.

Received 11 August; accepted 2 December 1980.

1. Oke, J. B. *Publ. astr. Soc. Pacif.* **81,** 11 (1969).
2. Chapman, C. & Gaffey, M. in *Asteroids* (ed. Gehrels, T.) (University of Arizona Press, Tucson, 1979).
3. Spinrad, H., Stauffer, J. & Newburn, R. L. Jr *Publ. astr. Soc. Pacif.* **91,** 707 (1979).
4. Zellner, B., Tedesco, E. & Degewij, J. *IAU Circ.* No. 3326 (1979).
5. Vsekhsvyatskii, S. K. in *Physical Characteristics of Comets* (IPST, Jerusalem, 1958).
6. Whipple, F. L. *Astr. J.* **85,** 305 (1980).
7. Cochran, A. L., Barker, E. S. & Cochran, W. D. *Bull. Am. astr. Soc.* **11,** 566 (1979).
8. Larson, S. M. *Astrophys. J.* **238,** L47 (1980).
9. Richter, N. *Astr. Nachr.* **277,** 12 (1948).
10. Tatum, J. B. & Gillespie, M. I. *Astrophys. J.* **218,** 569 (1977).
11. Haser, L. *Bull. Acad. R. Sci. Belg.* **43,** 740 (1957).
12. A'Hearn, M. F. & Cowan, J. J. *Astr. J.* **80,** 852 (1975).
13. Delsemme, A. H. in *The Study of Comets* (eds Donn, B., Mumma, M. Jackson, W. A'Hearn, M. & Harrington, R.) (NASA SP-393, 1976).
14. Delsemme, A. H. & Miller, D. C. *Planet. Space Sci.* **19,** 1229 (1971).
15. Sekanina, Z. & Miller, F. D. *Science* **179,** 565 (1973).
16. Newburn, R. L. Jr & Johnson, T. V. *Icarus* **35,** 360 (1978).

Stability of zeolites under electron irradiation and imaging of heavy cations in silicates

L. A. Bursill*, J. M. Thomas‡ & K. J. Rao†

Department of Physical Chemistry, University of Cambridge,
Lensfield Road, Cambridge CB2 1EP, UK

The aluminosilicate framework of synthetic 'zeolites' of type A (Na^+-form) and ZSM-5, as well as the Na^+-forms of faujasite (Fig. 1) in their so-called X- and Y-variants (fuller structural descriptions are given in refs 1 and 2) have recently been imaged at near-atomic scale by high-resolution electron microscopy (HREM)[3–5]. However, the as-prepared synthetic and naturally occurring zeolites, all of which contain alkali metal and alkaline-earth exhangeable cations and much constitutional, intracrystallite water, are converted within minutes into an amorphous state on examination in an electron beam of the intensity required to yield high-resolution images. The specimen has to withstand 10^6 electrons $Å^{-2}$ bombardment within a few minutes. To extract further information from HREM about the local atomic structure, and, in particular, to identify the siting of the various cations that may be accomodated in these catalytically important and shape-selective materials, methods have to be developed which retain the structural integrity of the zeolite for a reasonable time under electron irradiation *in vacuo*. It has been shown[3] that dehydration of the zeolite before observation significantly increases its lifetime in an electron beam. Here we show that the commercially important[4] zeolite-Y may be chemically stabilized against electron irradiation damage by the simple expedient of incorporating uranyl (UO_2^{2+}) ions into the structure beforehand. Beam sensitivities are decreased by factors of 10–100 in this way, but appropriate heat-treatment of the UO_2-exchanged zeolite is required to achieve this stability and hence to locate these heavy cations.

For the present study the UO_2^{2+} ion has two advantages. First, its propensity to form strong links with surrounding oxygen ligands: this along with the space-filling nature of this bulky ion should strengthen and rigidify the relatively fragile aluminosilicate framework of the zeolite. Second, the U atom ($Z = 92$) is a strong scatterer of electrons which should make it relatively easy to locate the centres of the linear UO_2^{2+} ions in high-resolution images.

Low-dose electron diffraction studies showed that straightforward cation-exchange of Na^+ by UO_2^{2+} in Na-Y and drying at ~90 °C does not significantly change the lifetime of the zeolite. Subsequent high-magnification imaging results in rather rapid amorphisation, as found previously for Na-A, -X and -Y. Simultaneously with this transformation to a zeolitic glass we observe (Fig. 2) growth of a thin surface film. Both the image, which was recorded for the optimum objective lens defocus condition ($\Delta f = -652$ Å) for a lens having spherical aberration coefficient $C_s = 1.2$ mm, and the corresponding optical transform (inset Fig. 2) show that interatomic spacings of 2.5–2.7 Å predominate. A survey of the various phases of uranium oxide and of uranium metal reveals that U–U separations, when projected normal to common low-index zone axes, should cluster around 3.5 Å for phases of the crystalline UO_3-type, around 2.2–3.3 Å for UO_2 and 2.5–2.9 Å for α-uranium metal. Clearly the surface film, the identity of which has not yet been fully characterized, shows a microcrystalline texture with short-range order extending over only 9–15 Å.

If, instead of the drying at ~90 °C, the UO_2^{2+}-exchanged zeolite is heated for extended periods *in vacuo* (typically 15 h at 410 °C), the beam resistance of the Y-zeolite is greatly enhanced. Thus the micrograph shown in Fig. 3 was obtained after the specimen had been subjected to at least 100 times the exposure previously required to amorphise Na–Y. The pseudo-hexagonal array of dark spots in Fig. 3*a* (spacings 17 and 16 Å) shows an essentially one-to-one correspondence with the positions of the 'decorated' super-cages in the [110] projection of zeolite-Y (that is faujasite—see Fig. 3*b*). Because the image was obtained using the condition of optimum objective lens defocus for structure imagining at ~5 Å resolution[6] we interpret the dark spots as representing the projected charge density of UO_2^{2+} ions occupying the zeolite cages.

This specimen, the IR spectrum[7,8] of which indicated that the UO_2^{2+} ions were still intact and that the OH frequencies were greatly reduced in intensity after heating to 410 °C, could be given short bursts of extremely high electron irradiation (by

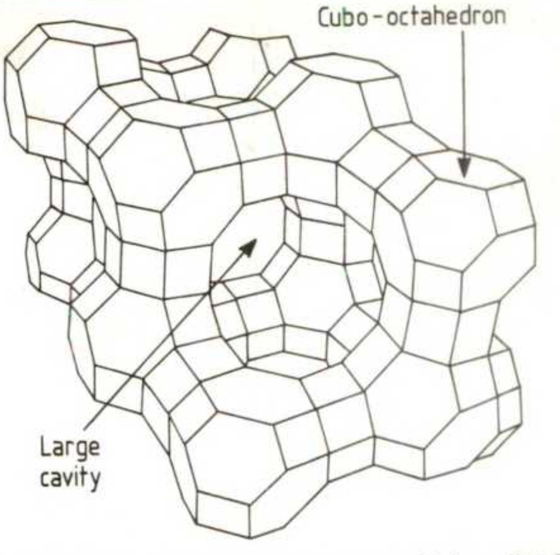


Fig. 1 Faujasite, the parent structure on which zeolites-X and -Y are based (Si:Al ratios differ in X and Y and they differ also from the ratio of 2.48 in idealized faujasite).

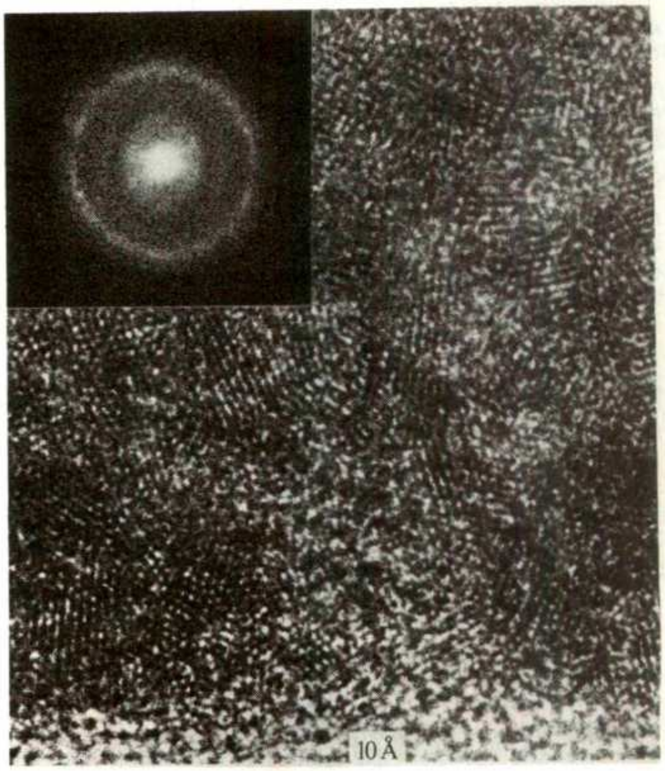


Fig. 2 High-resolution image with optical diffractogram (inset) of microdomain texture produced on the surface of hydrated uranyl-exchanged zeolite-Y on amorphisation in a 200-keV electron beam (see text).

Permanent addresses: *School of Physics, The University of Melbourne, Parkville 3052, Australia; † Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560012, India.
‡ To whom correspondence should be addressed.

0028-0836/81/030157—02$01.00

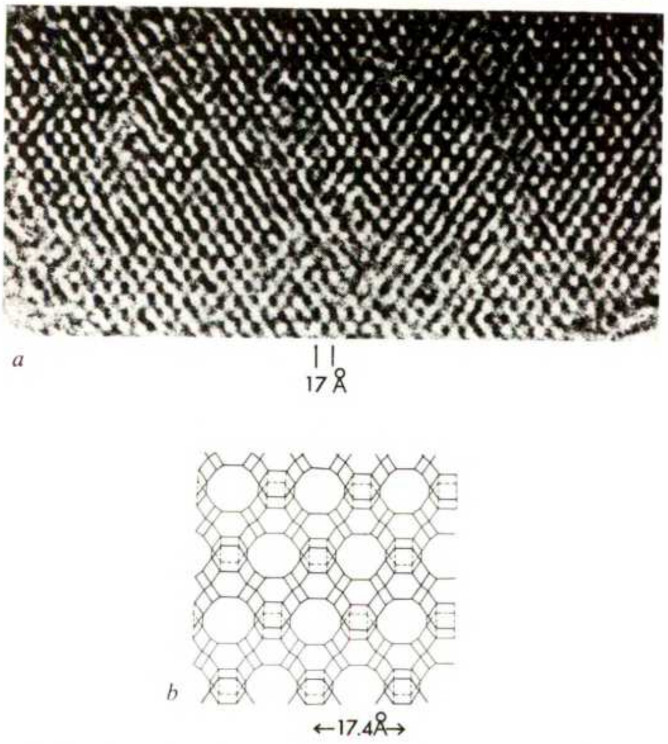


Fig. 3 *a*, High-resolution image along [110] showing location of uranyl ions inside a dehydrated uranyl-exchanged zeolite-Y. This crystal survived a dose of ≈100 times that experienced by the specimen shown in Fig. 2. *b*, Schematized projection, along [110], of faujasite (zeolite-Y) showing that the large cavities are 17.4 and 15.1 Å apart. The dark spots in *a* are, therefore, identified as the cavities decorated with UO_2^{2+} ions (see text).

removing condenser apertures and focusing the illumination) without any apparent change in structural integrity. Excessive beam currents did, however, lead to the growth of electron-opaque spherical crystallites (~300 Å diameter) within a matrix of amorphized zeolite. This confirms that the Y-zeolite contained, as suggested by Fig. 3*a*, an initially homogeneous distribution of incorporated UO_2^{2+} ions. The mobility of the specimen, when subjected to excessive beam currents, suggests that the temperatures reach well beyond 1,000 °C.

Although the precise environment of the UO_2^{2+} ion after high-temperature treatment (410 °C) and hence the detailed reasons for the enhanced beam stability are not yet fully identified, it is plausible to suppose that residual water has a key role (see also ref. 1). If water molecules are present in appreciable quantities (either as hydration shells or bound to cage walls) a ready mechanism exists for the exsolution of the UO_2^{2+} ions. Thus, at ~90 °C, the substitution of $2H_3O^+$ for UO_2^{2+} would be possible. At 410 °C, as the IR results reveal, this substitution cannot occur: the UO_2^{2+} ions are firmly chemically bound to, and thereby strengthen, the aluminosilicate framework.

The possibility of using UO_2^{2+} ions to decorate or label cages, cavities, tunnels and interlamellar spaces in other crystalline or non-crystalline aluminosilicates, as well as of charting the connectivity of the voids in, and the diffusion of heavy ions through, such materials is being investigated.

There is considerable technological interest in the stability and environment adopted by uranium in various host matrices. This stems from the need to sequester actinide waste products effectively in environments that prevent migration of the radioactive ions into the biosphere. During this work, we have incidentallly discovered that both uranium glasses and uranyl ions embedded in 'dehydrated' zeolites are remarkably stable to quite intense 200 keV electron bombardment.

We thank the SRC, and the University of Melbourne for support, and Drs S. Vasudevan, G. R. Millward and D. A. Jefferson for cooperation.

Received 14 October; accepted 6 November 1980.

1. Meier, W. M. & Olsen, D. H. *Atlas of Zeolite Structure Types* (Juris Druck, Zurich, 1978).
2. Barrer, R. M. *Zeolites and Clay Minerals as Sorbents and Molecular Sieves* (Academic, London, 1978).
3. Bursill, L. A., Lodge, E. A. & Thomas, J. M. *Nature* **286**, 111–113 (1980).
4. Thomas, J. M., Millward, G. R. & Bursill, L. A. *Phil. Trans. R. Soc.* (in the press).
5. Thomas, J. M. & Bursill, L. A. *Angew. Chem.* **92**, 755–756 (1980).
6. Thomas, J. M. *New Scient.* **87**, 580–584 (1980).
7. Lynch, D. F., Moodie, A. F. & O'Keefe, M. A. *Acta crystallogr.* **A31**, 300 (1975).
8. Deane, A. M. *J. Inorg. nucl. Chem.* **21**, 238 (1961).
9. Terada, M. & Tsubri, M. *Bull chem. Soc. Jap.* **37**, 1080 (1964).

Photochemical cleavage of water by photocatalysis

Enrico Borgarello*, John Kiwi*, Ezio Pelizzetti†, Mario Visca‡ & Michael Grätzel*

* Institut de Chimie Physique, Ecole Polytechnique Fédérale, Lausanne, Switzerland
† Istituto di Chimica Analitica, Università di Torino, Italy
‡ Centro Richerche SIBIT (Montedison), Spinetta Marengo, Italy

A bifunctional redox catalyst, composed of Pt and RuO_2 co-deposited on a colloidal TiO_2 carrier, is a highly potent mediator for water decomposition by visible light[1]. The system contains apart from the sensitizer ($Ru(bipy)_3^{2+}$) an electron relay—methylviologen. The latter is reduced on light excitation, and the photoreaction is coupled with catalytic steps[2] generating H_2 and O_2 from water. To rationalize the surprisingly high efficiency of this photoredox system, we proposed a mechanism involving species adsorbed at the TiO_2 surface. This led us to explore sensitizers which through suitable functionalization show an enhanced affinity for adsorption at the particle–water interface. We describe here the performance of electron relay-free systems capable of efficiently decomposing water into H_2 and O_2 under visible light illumination. A bifunctional redox catalyst composed of RuO_2 and Pt co-supported on colloidal TiO_2 particles is used. The only other component present is a sensitizer. Amphiphilic surfactant derivatives of $Ru(bipy)_3^{2+}$ exhibit extremely high activity in promoting the water cleavage process. Adsorption of the sensitizer at the TiO_2 particle–water interface and electron ejection into the TiO_2 conduction band are evoked to explain the observations. Exposure to UV radiation leads to efficient water cleavage in the absence of sensitizer.

The successful operation of this device depends on the proper preparation of the catalyst. In our experiments 100 mg of H_2PtCl_6 was added to 250 ml of water and stirred for 30 min without heating. 85 ml of a 1% sodium citrate solution was added and the mixture refluxed for several hours until the colour changed from yellow to brown. At this time the characteristic absorption of $PtCl_6^{2-}$ has disappeared, the spectral features in the visible and UV being typical of those of an ultrafine Pt sol[3]. The absorbance at 450 nm is 0.55 (pathlength 1 cm). After cooling, the solution was stirred with Amberlite MB 3 exchange resin to remove excess citrate until the conductivity decreased below 5 μs cm^{-1}. The solution was filtered through a 0.45 μm Millipore filter and a clear liquid was obtained with an absorbance of 0.50 at 450 nm. 12 mg of TiO_2 was added to 12 ml of this solution and exposed for several minutes to ultrasound in a water bath.

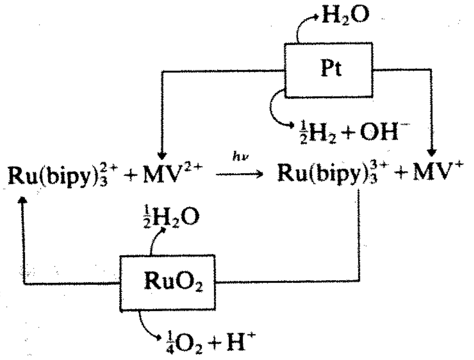
The TiO_2 particles have an anatase structure and were prepared from $TiOSO_4$ by acid hydrolysis. The surface area is 200–240 m^2 per g and the hydrodynamic radius of the particles 470 Å. They are also doped by 0.1% RuO_2 and by 0.4% Nb_2O_5.

Illuminations were carried out with an XBO-450 W Xe lamp fitted with a water jacket to remove IR irradiation. The gas produced was analysed by gas chromatography. A Gow Mac system, carbosieve column (35 °C) and Ar (for H_2) or He (for O_2) carrier gas was used. The volume of the solution was invariably 24 ml and contained 500 mg l^{-1} TiO_2 as well as 40 mg l^{-1} Pt; it was contained in a Pyrex flask fitted with a septum. The solution was deaerated before illumination by flushing with Ar or He.

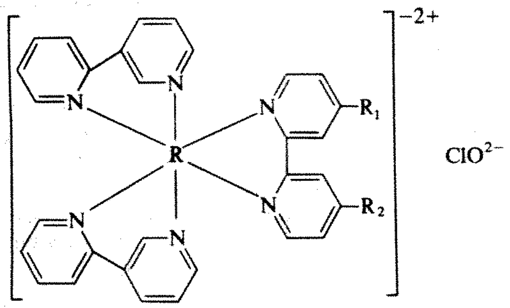
0028–0836/81/030158—03$01.00

The colloidal $Pt/TiO_2/RuO_2$ particles were first irradiated without the sensitizer. If no cutoff filter is used to remove UV irradiation (in this case transmission is restricted by the Pyrex flask) production of H_2 and O_2 is observed through direct band gap excitation of the semiconductor particles. After 10 h exposure to light 3.4 ml of H_2 was evolved. Oxygen was formed in stoichiometric proportions. Stirring of the solution in the dark for 10–20 h results in H_2/O_2 recombination. This reaction can be avoided by removing the gas mixture after photolysis with a stream of N_2. On renewed exposure to light, the water splitting continues at the initial rate, 0.34 ml H_2 h^{-1}. This cycle can be repeated many times. H_2 and O_2 generation is completely suppressed if a 400-nm cutoff filter is placed in the light beam. These results show that the particles can split water into hydrogen and oxygen after band gap excitation. The yields are surprisingly high showing efficient electron–hole separation. Water splitting on TiO_2 and $SrTiO_3$ powders has been observed before[4–9], but with much lower efficiency.

The $Pt/TiO_2/RuO_2$ particles can also be used as catalysts for water splitting by visible light in the presence of a suitable sensitizer and methylviologen as an electron relay[1],



The present study shows that this system can be made operative even in the absence of the electron relay. The sensitizers used were ruthenium trisbipyridyl complexes of the structure:



where

$R_1 = R_2 = H$: $Ru(bipy)_3^{2+}$
$R_1 = R_2 = n$-hexadecyl: $Ru(bipy)_3^{2+} \cdot 2C_{16}$
$R_1 = CH_3$, $R_2 = n$-octadecyl: $Ru(bipy)_3^{2+} \cdot C_{18}$
$R_1 = CH_3$, $R_2 = n$-doceyl: $Ru(bipy)_3^{2+} \cdot C_{12}$

Data from visible light illumination of solutions containing the sensitizer and the bifunctional redox catalyst are displayed in Fig. 1. The amount of hydrogen generated is compared for the four different sensitizers. Oxygen remained in stochiometric ratio to the H_2 produced. These solutions contained no buffer and were adjusted to a pH of 4.5 by addition of HCl. No pH changes could be detected even after prolonged irradiation. In Fig. 1 the lowest H_2 production rate is observed with the simple $Ru(bipy)_3^{2+}$ complex. Introduction of a long alkyl chain leads to a drastic improvement in the H_2 yields. Of the surfactant derivatives available[10], the mono C-12 substituted complex exhibits optimal efficiency. After a short induction period, the H_2 generation rate was established at 1.5 ml h^{-1}. In view of the small concentration of sensitizer ($\sim 5 \times 10^{-5}$ M), this rate is surprisingly high. At this concentration, light absorption is incomplete even at λ_{max} of the sensitizer. As the solutions are slightly turbid scattering of light cannot be neglected making the determination of quantum yield difficult. However, from a comparison with a non-regenerating H_2-producing system[11] ($\phi(H_2) = 0.13$) used in identical conditions, the quantum yield of the H_2 production with the C-12 substituted Ru-complex is derived as ~5%. This may be further improved by optimization of the sensitizer, catalyst and pH condition.

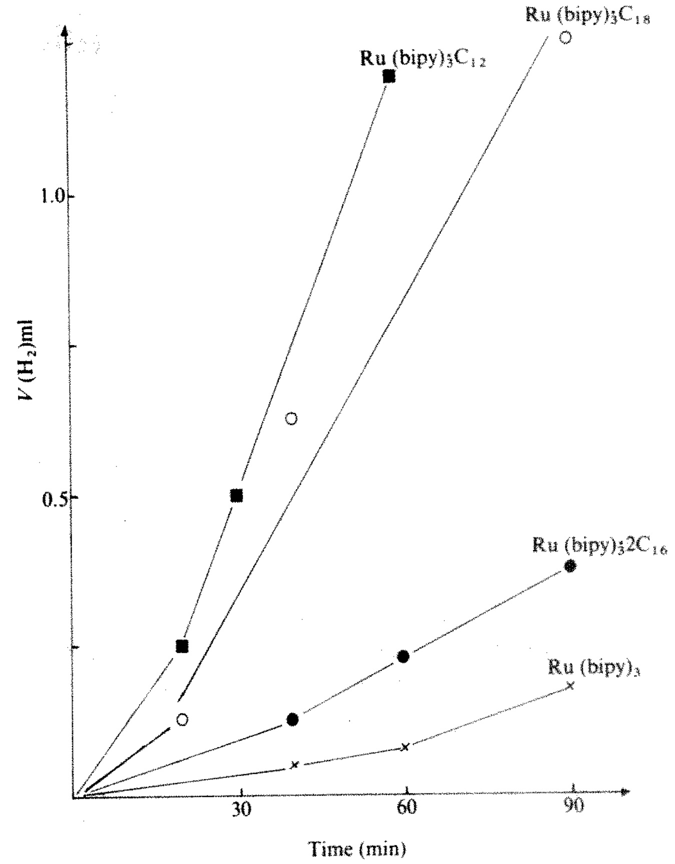


Fig. 1 Hydrogen yields from visible light (>400 nm) photolysis of solutions containing 500 mg per l TiO_2/RuO_2 (0.1%) doped with 40 mg per l Pt, $pH = 4.5$ adjusted with HCl. Sensitizer concentration:

$Ru(bipy)_3^{2+} = 2 \times 10^{-4}$ M;

$$\left.\begin{array}{l} Ru(bipy)_3^{2+} \cdot 2C_{16} \\ Ru(bipy)_3^{2+} \cdot C_{18} \end{array}\right\} = 7 \times 10^{-5}\ M;$$

$Ru(bipy)_3^{2+} \cdot C_{12} = 5 \times 10^{-5}$ M.

The H_2 production rate does not decrease even after 50 h of irradiation time, the turnover for the Ru complex being >1,000.

These observations may be rationalized in terms of the scheme shown in Fig. 2. The effect of functionalization of the $Ru(bipy)_3^{2+}$ on the H_2 yields confirms that the photoredox processes involve adsorbed species. Excitation of the sensitizer is followed by charge injection into the conduction band of TiO_2. The electron is readily channelled to Pt sites where H_2 evolution occurs. A small Schottky potential barrier at the Pt/TiO_2 junction may assist the electron movement. The back conversion of oxidized sensitizer into its original form is coupled to oxygen formation through RuO_2 which is an excellent electrocatalyst for water oxidation.

The effects observed when the TiO_2 particles are directly excited by band gap irradiation have been well understood since Honda's[12] work on the cleavage of water by UV-ray irradiation of TiO_2 macroelectrodes. The surprisingly high quantum yields of the UV process show that electron–hole separation in the particles is efficient and probably assisted by local fields.

Cleavage of water is possible by four quanta of visible light in an electron relay-free system containing simply a sensitizer and colloidal redox catalyst. A key to achieve high efficiencies is the

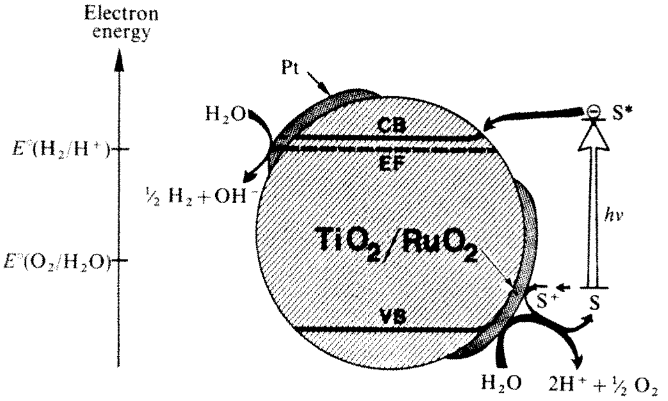


Fig. 2 Processes involved in the photodecomposition of water.

use of surfactant redox chromophores[13–15] which are adsorbed at the surface of the catalytic particles. The water cleavage on such functional microspheres can thus occur without participation of diffusional processes. These results should encourage research into photochemical conversion of solar energy[16–20].

We thank Sir George Porter for discussions on electron relay-free systems and Dr W. F. Sasse for samples of the surfactant ruthenium complexes. This work was supported by the Swiss NSF, Engelhard Industries, and CIBA-GEIGY, Basel.

Received 11 September; accepted 4 November 1980.

1. Kiwi, J., Borgarello, E., Pelizzetti, E., Visca, M. & Grätzel, M. *Angew. Chem.* **92,** 663 (1980).
2. Kalyanasundaram, K. & Grätzel, M. *Angew. Chem. int. edn Engl.* **18,** 701 (1979).
3. Turkevich, J., Aika, K., Ban, L. L., Okura, I. & Numba, S. *J. Res. Inst. Catal. Hokkaido Univ.* **24,** 54 (1976).
4. Bulatov, A. V., & Khidekel, M. L. *Izv. Akad. Nauk SSSR Ser. Khim.* 1902 (1976).
5. Schrauzer, G. N. & Gath, T. D. *J. Am. chem. Soc.* **99,** 7189 (1977).
6. Van Damme, H. & Hall, W. K. *J. Am. chem. Soc.* **101,** 4373 (1979).
7. Sato, S. & White, J. M. *Chem. phys. Lett.* **72,** 83 (1980).
8. Kawai, T. & Sakata, T. *Chem. phys. Lett.* **72,** 87 (1980).
9. Domen, K., Naito, S., Sorna, M., Onishi, T. & Tamaru, K. *JCS Chem. Commun.* 543 (1980).
10. Johansen, O., Kowala, C., Mau, A. W.-H. & Sasse, W. H. F. *Aust. J. Chem.* **32,** 1453 (1979).
11. Kiwi, J. & Grätzel, M. *Nature* **281,** 657 (1979).
12. Fujishima, A. & Honda, K. *Nature* **238,** 37 (1972).
13. Clark, W. D. K. & Sutin, N. *J. Am. chem. Soc.* **99,** 4676 (1977).
14. Hamnet, A., Dare-Edwards, M. P., Wright, R. D., Seddon, K. R. & Goodenough, J. B. *J. phys. Chem.* **83,** 3280 (1979).
15. Mackor, A. & Schoonman, J. *Rec. J. Neth. chem. Soc.* **99,** 71 (1980).
16. Balzani, V., Moggi, L., Manfrin, M. F., Bolletta, F. & Gleria, M. *Science* **189,** 852 (1975).
17. Calvin, M. *Photochem. Photobiol.* **23,** 425 (1976).
18. Porter, G. & Archer, M. D. *Interdisc. Sci. Rev.* **1,** 119 (1976).
19. Bolton, J. R. *Science* **202,** 105 (1978).
20. Schumacher, E. *Chimia* **32,** 193 (1978).

Crustal influence in the generation of continental flood basalts

R. W. Carlson*, G. W. Lugmair & J. D. Macdougall

Scripps Institution of Oceanography A-020, University of California, San Diego, La Jolla, California 92093

The suggestion that primordial undifferentiated material may exist in the Earth's mantle has recently been revived on the strength of Nd isotope data for two types of young continental rocks—flood basalts[1] and kimberlites[2]. The limited published data show a clustering of Nd isotopic compositions close to those for meteorites with chondritic relative rare-earth (REE) abundance. In contrast, we now present data for samples from the Columbia flood basalt province of the northwestern United States which show large isotopic variability suggestive of mixing processes acting after the separation of the primary magmas from their mantle source.

The Columbia province contains ~200,000 km^3 of predominantly basaltic rocks which were fed from fissure eruptions 6–17 Myr ago[3]. The province is in the volcanically and tectonically active Pacific north-west of the US, to the east of the present locus of Cascade andesitic arc volcanism. This area is the most extensively studied of the major flood basalt provinces[3–5]. Five main stratigraphic units in the basalts have been identified and in rough stratigraphic order, are named: Imnaha, Picture Gorge, Grande Ronde, Wanapum and Saddle Mountains. The field relations and major element composition of these units are detailed elsewhere[3].

We have analysed the Nd and Sr isotopic composition of samples from each of the stratigraphic units, and from a series of unclassified Columbia basalts of variable composition which overlie the Grande Ronde in northeastern Oregon[6] and also mafic lavas from various localities in northern California. Collection localities and major and trace element composition of these samples are given elsewhere[7]. The Sr data measured here augment studies[5] of the Columbia region, and demonstrate how combined Nd and Sr isotopic analysis aids the understanding of the origin of complex chemical and isotopic variations found in these lavas.

Isotopic results are shown in Fig. 1. Analytical techniques have been described elsewhere[7,8]. All Nd data are fractionation corrected to ^{148}NdO/^{144}NdO = 0.242436 (corresponding to ^{146}Nd/^{144}Nd = 0.7219). Sr values are fractionation corrected to ^{86}Sr/^{88}Sr = 0.1194 and are reported relative to a value of ^{87}Sr/^{86}Sr = 0.71014 for NBS 987 Sr.

Figure 1 shows the large isotopic variability within the basalts of the Columbia group. Particularly important is the good correlation between Nd and Sr isotopic variation. Note the general correlation between isotopic composition and stratigraphic position as the variability extends in the order, Picture Gorge ($\varepsilon_{JUV} = +7.7$ to $+7.4$), Imnaha (7.8 to 5.9), Grande Ronde (+4.0 to +0.3), Wanapum (+1.2 to +0.6) and Saddle Mountains (−3.6 to −10.8). In particular, the basalts of the Grande Ronde group, which represent the largest volume of the Columbia sequence, exhibit ε_{JUV} values ranging from +4.0 to +0.3, which correlates roughly with stratigraphic position within this group. The volume weighted average Nd isotopic composition is thus resolvably displaced to values of $\varepsilon_{JUV} > 0$. However, the high degree of correlation of Nd and Sr isotopic variations suggests that this range may be the result of mixing between at least two components: primary magma with $\varepsilon_{JUV} \geq 8$, ^{87}Sr/^{86}Sr ≤ 0.7033, and a source of negative ε_{JUV} and high ^{87}Sr/^{86}Sr. At least two sources of negative ε_{JUV} and high ^{87}Sr/^{86}Sr can be proposed: (1) old, metasomatically enriched, mantle[9,10]; (2) materials derived from the continental crust underlying the province. Verification of the former possibility is hindered by the lack of knowledge of the degree and time scales over which heterogeneities are formed and maintained in the mantle. Similar Nd and Sr covariation, to that displayed in Fig. 1, have been reported for volcanic rocks from Patagonia[11] and Italy[12]. In the latter case, the co-variation was interpreted to be the result of mantle heterogeneity rather than crustal contamination primarily due to a positive correlation between Sr concentration and ^{87}Sr/^{86}Sr in these lavas. This argument cannot be applied to the Columbia basalts because they exhibit, in general, poor correlations between Sr and ^{87}Sr/^{86}Sr. In the only group that shows good Sr versus ^{87}Sr covariation, the Saddle Mountains group, Sr concentration decreases with increasing ^{87}Sr/^{86}Sr. Support for a chemically and isotopically heterogeneous source for the Columbia lavas comes from the fact that major and trace element compositions of these lavas cannot be simply related by mixing between two end-member components[4,5]. For example, the Saddle Mountains group contains members which are relatively primitive chemically, and others which are highly fractionated, although these all have $\varepsilon_{JUV} \leq -3.6$ and ^{87}Sr/^{86}Sr ≥ 0.7075. However, considering the complexity of magma storage and transport in the upper mantle and crust, it is possible to construct a model which explains much of the elemental and isotopic diversity found for the Columbia lavas, assuming that they were all generated from a single homogeneous source.

The inset to Fig. 1 includes Nd and Sr isotopic data for a xenolith recovered from a flow of the Wanapum group. This xenolith was ~3 cm in diameter and consisted primarily of quartz with minor cristobalite. Its major element composition

---

* Present address: Department of Terrestrial Magnetism, 5241 Broad Branch Road, NW Washington DC 20015.

0028-0836/81/030160—03$01.00

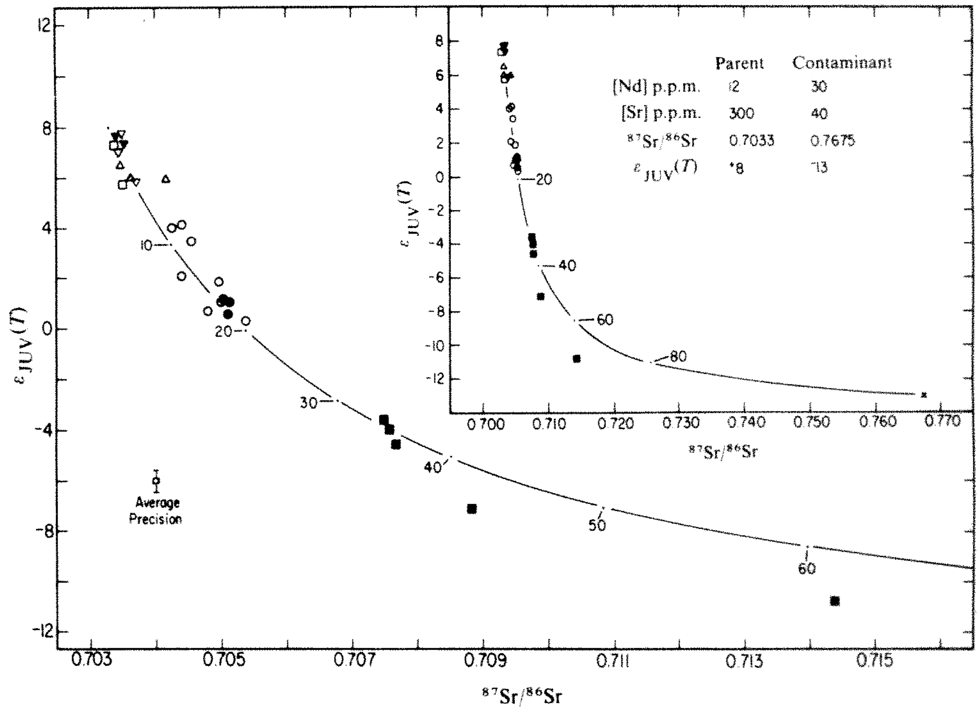


Fig. 1 Nd–Sr isotopic variation of Columbia River and some northern California basalts. Data for a metasedimentary crustal xenolith are included in the inset (×). The line is a calculated mixing curve between end-member compositions given in the inset. Numbers on the curve are the percentage of 'contaminant' in the mixture. The deviation from the curve of the two points with large negative ε_{JUV} suggests that these two basalts may have assimilated material with lower $^{143}Nd/^{144}Nd$ and/or $^{87}Sr/^{86}Sr$ than measured for the xenolith. All data have been corrected for isotopic evolution since the time of eruption. ▼, Picture Gorge; ▽, Imnaha; ○, Grande Ronde; ●, Wanapum; ■, Saddle Mountains; □, NE Oregon; △, North California. ε_{JUV} is the relative difference of the samples $^{143}Nd/^{144}Nd$ (in parts per 10,000) from that measured for the eucrite Juvinas ($^{143}Nd/^{144}Nd = 0.512636$).

and texture suggest a metasedimentary crustal protolith. This xenolith has very high $^{87}Sr/^{86}Sr$ (0.768) and negative ε_{JUV}. Sm–Nd and Rb–Sr model ages for the xenolith, calculated in the T_{ICE} manner[8], are 1,320 and 1,420 Myr, respectively. The near agreement of these two ages suggests both that little Rb–Sr or Sm–Nd fractionation occurred during the entrapment of the xenolith in the lava, and that this may represent the formation age of the protolith material. If true, this xenolith may be the first identified example of a Precambrian crustal section existing beneath the Columbia Plateau.

While both the major element and oxygen isotopic composition of this particular xenolith argue against its being the specific crustal end member involved, it can be used to illustrate what effect assimilation of crustal materials would have on the Sr and Nd isotopic composition of the Columbia lavas. For example, Fig. 1 shows a calculated mixing curve between end-member compositions given in the inset. The 'parent' end member has the composition of the primitive lavas of the Picture Gorge sequence. The 'contaminant', with identical Nd and Sr isotopic composition and Sr concentration as the measured xenolith, requires an ~78% higher Nd concentration than found for the xenolith to give the curve in Fig. 1. The calculated curve lies within analytical uncertainty of all but two basalt data points. Similar mixing calculations using a crustal end member with less radiogenic Nd and Sr isotopic composition and a higher Sr/Nd would produce a mixing curve with less inflection than that shown and hence provide a better fit to the basalt data. Although the exact crustal end member is not defined by present data, the coherence of the basalt Nd and Sr isotopic data about a single curve strongly suggests that mixing between primary magma and materials derived from the continental crust played a major part in controlling the isotopic variability of the lavas.

Using the end-member compositions given in Fig. 1, magmas of the Grande Ronde group may have assimilated between 10 and 20% of their Sr and Nd contents from the continental crust. Considerations of the thermal effects of basalt intrusion into the continental crust[13] also suggest that "basalt sheets at 10 km depth, for example, can develop partially molten zones of country rock equal to 10–20% of their own thickness". This implies that the ε_{JUV} values which approach zero for the largest volume of the Columbia group, the Grande Ronde, may reflect a most probable mass balance between primary magma and assimilated crust due to the degree of heat transfer possible from the Grande Ronde lavas to the crust. This hypothesis could also be extrapolated to the other occurrences of continental flood basalts, but is, of course, subject to the elemental and isotopic composition of the crustal end member which may have been involved in each province.

Although the Nd and Sr isotopic variation of the Columbia lavas can be explained by mixing between two components, major and trace elements do not show good co-variation among the groups defined by the isotopic data. However, when controlled by mixing, the resultant isotopic composition of the mixture is determined only by the mass fractions of the various components. This is not necessarily true for the major and trace element composition of the mixture. If processes such as fractional crystallization and various degrees of mixing between highly fractionated magmas and newly injected, rather primitive, magmas occurred simultaneously with crustal assimilation, major and trace element composition of the resulting magmas would be controlled largely by these processes rather than assimilative mixing. Furthermore, the passage of these lavas through the crust may have had a considerable effect on the chemical composition of the residual materials left behind in the crust. In a volcanic province where eruption occurred at a rather rapid rate over a considerable span of time (11 Myr for Columbia[3]) magmas might be expected to accumulate preferentially and be transported in paths and storage reservoirs created by previous magma batches. Therefore, later magmas would not encounter the original 'sialic' crust but would have to interact with the more mafic material left behind by previous magma batches. This could account for the very high Fe and Ti and relatively low Si concentrations of the Wanapum group compared with the lavas of the Grande Ronde sequence, although these two groups have overlapping ranges in Nd and Sr isotopic composition.

A qualitative model for both the compositional and isotopic variation in the Columbia basalts assuming that they were derived from a homogeneous source region, would require: (1) different degrees or conditions of partial melting in the mantle; (2) transport and storage of primary magmas in crustal storage reservoirs where assimilation of Precambrian crustal material of varying composition, fractional crystallization in response to assimilation, and mixing between dense, Fe rich, highly fractionated liquids and newly injected, rather low density, unfractionated magmas occurred.

Picture Gorge magmas may have escaped the effect of this crustal contamination by their eruption from vents occurring well to the south-west of the majority of dyke swarms in the Columbia province. This hypothesis requires complex modelling and a better understanding of what really occurs when a basic magma is injected into the lower continental crust before quantitative modelling is possible. However, the very good correlation between Nd and Sr isotopic variation in these lavas and the demonstration that Precambrian continental crust exists under the province, strongly supports the viability of crustal contamination as an explanation for the chemical characteristics of the Columbia group.

If these lavas were derived from a homogeneous source and later modified by a series of complex geochemical interactions, then the parental mantle of the Columbia lavas had an isotopic composition of $^{87}Sr/^{86}Sr \leq 0.7033$ and $\varepsilon_{JUV} \geq +8$. These values are very similar to those measured for interoceanic island arcs and marginal basins of the Western Pacific. Whether these results can be extrapolated to other flood basalt provinces is unknown. However, the suggestion of a primordial (that is, $\varepsilon_{JUV} = 0$) source for the Columbia lavas[1] is clearly in error. The present results for the Columbia group do not exclude or confirm the existence of primordial material in the mantle. Rather they indicate the importance of the physical, chemical, and thermal phenomena which would accompany the transport and eruption of large volumes of mafic lavas, such as those of the Columbia province, through an ~30 km thick section of continental crust in a relatively short period of time.

We thank J. Holbrook for performing chemical separations, and W. H. Taubeneck, D. A. Swanson, T. L. Wright, E. Mullen and R. Stern for providing some of the samples. M. Menzies and M. Tatsumoto provided constructive reviews. This work was supported by NASA grant NGL05-009-150 and NSF grant EAR78-24060.

Received 15 August; accepted 31 October 1980.

1. DePaolo, D. J. & Wasserburg, G. J. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3056–3060 (1979).
2. Basu, A. R. & Tatsumoto, M. *Science* **205,** 398–401 (1979).
3. Swanson, D. A., Wright, T. L., Hooper, P. R. & Bentley, R. D. *Bull. U.S. geol. Surv.* 457G, (1979).
4. Wright, T. L., Grolier, M. J. & Swanson, D. A. *Bull. geol. Soc. Am.* **84,** 371–386 (1973).
5. McDougall, I. *Bull. geol. Soc. Am.* **87,** 777–792 (1976).
6. Taubeneck, W. H., *Pacif. Div. AAAS A. Meet.* (1979).
7. Carlson, R. W. thesis, Univ. California (1980).
8. Lugmair, G. W., Scheinin, N. B. & Marti, K. *Proc. 6th Lunar Sci. Conf.* 1419–1429 (1975).
9. Menzies, M. & Murthy, V. R. *Nature* **283,** 634–636 (1980).
10. Stoch, H.-G., Carlson, R. W. & Lugmair, G. W. *Earth planet. Sci. Lett.* **47,** 263–271 (1980).
11. Hawkesworth, C. J. *et al. Earth planet. Sci. Lett.* **42,** 45–57 (1979).
12. Hawkesworth, C. J. & Vollmer, R. *Contr. Miner. Petrol.* **69,** 151 (1979).
13. Patchett, P. J. *Nature* **283,** 559–561 (1980).

U series and amino acid dates from Jersey

D. H. Keen*, R. S. Harmon† & J. T. Andrews‡

* Department of Geography, Coventry (Lanchester) Polytechnic, Coventry CV1 5FB, UK
† Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 OQU, UK
‡ INSTAAR and Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309

Understanding of the raised beach and other interglacial sequences in the southern part of the British Isles has been improved through use of the amino acid racemization technique. This method has permitted correlation of these isolated and fragmentary marine deposits[1,2] which, except for the Portland site, are located around the shores of the Bristol Channel[1]. By applying this technique with the U-series method, we have obtained a geochronometric date from the '8-m' raised beach of the Belle Hougue Cave, Jersey, Channel Islands. Our results suggest that the 8-m raised beach of the Channel should be correlated with isotope stage 5, and that raised beaches above this altitude in the area may pre-date this stage, and thus also the Ipswichian interglacial which forms part of it.

Although there are many raised beaches around the coasts of the western English Channel[3–5] the non-calcareous nature of the bedrock geology of the area means that few of these are fossiliferous. The raised beach of the Belle Hougue Cave, although on a base of rhyolitic agglomerate of the Jersey Volcanic series, is strongly cemented by interstitial calcium carbonate and in places contains discrete travertine layers. In the matrix of the beach are the remains of nine species of Mollusca and mammalian bones[6]. Specimens of *Patella vulgata* Linné and fragments of travertine were collected from the beach for analysis (Fig. 2), the shells for the determination of amino acid racemization ratios and the travertine for U-series dating. The resulting D-allo:L-iso ratios and $^{230}Th/^{234}U$ dates place an absolute age on the Belle Hougue beach and provide reasonable grounds for correlation with other more northern raised beach sites.

The Belle Hougue Cave lies on the north coast of Jersey (Fig. 1). The beach gravel occurs ~25 m in from the cave entrance and has a height range between 4 m and 8 m above mean sea level (Fig. 2). The gravel contains abundant molluscan remains. Three shells of *P. vulgata* were collected from the main shell concentration in the beach which lies between 30 and 60 cm below the top of the deposit (Fig. 2). The fragmented shells were heavily cemented by $CaCO_3$ and required careful cleaning before being prepared and analysed for amino acid ratios[2]. The total ratio of D-alloisoleucine to L-isoleucine is given in Table 1. All of the analyses are similar giving a mean D-allo:L-iso value of 0.12±0.02, except AAL-1219b which has a significantly higher value. This anomalous value may be related to the high degree of cementation of this sample.

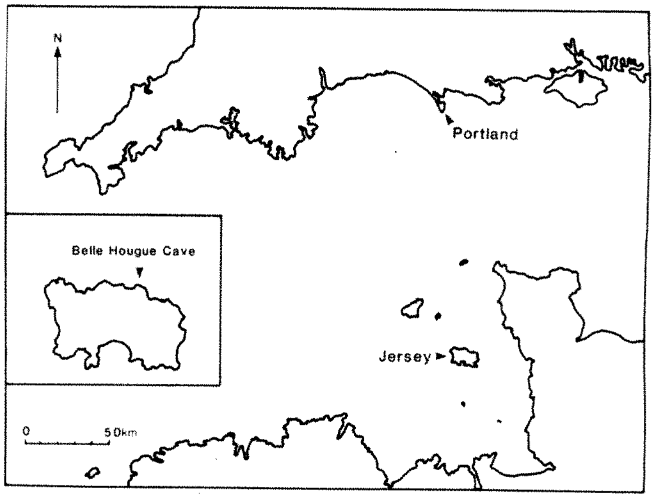


Fig. 1 Location diagram of the western Channel showing the positions of Jersey and Portland. Inset, the site of the Belle Hougue Cave.

The travertine was collected from a layer largely free of pebbles and shell, ~20 cm above the mollusc sample point (Fig. 2). A sample of 200 g was collected and subsplit for duplicate U-series analysis following the technique of Harmon *et al.*[7]. Analytical data and the U-series age of the sample are given in Table 1. The age of ~121 kyr BP indicated by the two runs has a slightly greater than normal uncertainty because of the presence of some detrital Th, and the fact that Th yields were low (<10%) due to the slightly contaminated nature of the travertine. The $^{234}U/^{238}U$ ratio, although perhaps rather high, is typical of the ratios obtained from secondary carbonates. Nevertheless, an age within isotopic stage 5, the last interglacial in the marine oxygen-isotopic record[8], is clear and a stage 5e correlation is suggested.

A stage 5e age for the beach is supported on palaeontological grounds. One of the species of Mollusca found in the beach gravel is *Astralium rugosum* (Linné) which has a present northern limit near La Rochelle, some 340 km south of Jersey. The occurrence of *A. rugosum* probably indicates a sea surface temperature for the Belle Hougue beach of 3–4 °C above that of the present temperature of the English Channel, a value

Table 1 Amino acid racemization analysis

Sample no.	D-allo:L-iso (combined)
AAL-1219 a	0.114
b	0.17
c	0.112
d	0.13
e	0.106
f	0.136
Mean of six samples	0.123±0.024

U-series analysis

Sample no.	U(conc) (p.p.m.)	[^{230}Th/^{234}U]	[^{234}U/^{238}U]	[^{230}Th/^{232}Th]
79050-1/2	0.46	0.69±0.04	1.56±0.02	16

Age (kyr BP) 121^{+14}_{-12}

compatible with that expected from the higher temperatures generally suggested for the last interglacial in both marine[8] and terrestrial records[9].

Although a broad stage 5e age for the Belle Hougue raised beach is clear, in terms of wider correlations, several problems occur. The closest agreement of the amino acid values with those noted by Andrews *et al.*[1] is with their group 2 sites at Portland and in the Burtle Beds, which give amino acid values between 0.11 and 0.15. However, these sites both relate to sea levels about 15 m above OD[10,11] and Portland contains a fauna rather cooler in aspect than that of the present Channel[12]. On faunal grounds, therefore, correlation from Belle Hougue to these sites further north is not exact. Slightly different post-depositional environmental conditions may have occurred at Belle Hougue giving higher amino acid values than would be the case further north where lower temperatures prevailed for the period after the abandonment of the raised beaches. The present mean annual temperature at Portland is 10.7 °C, while in Jersey (St Helier) it is 11.9 °C. A temperature difference of at least this magnitude has probably occurred throughout the past 100 kyr.

Despite the problem of the correlation of the amino acid and sea levels, the absolute age for a raised beach at ~7.5 m above mean sea level supports Bowen's contention[13] that beaches at this general level in western Channel are of Ipswichian (stage 5) age, and not dated to an earlier interglacial as suggested by Stephens[14]. Outside the Channel similar sea level stands ~5–10 m above present sea level have been dated by ^{230}Th/^{234}U to ~120 kyr.

In Mallorca, Butzer's[15] cycle YI is dated to 120 kyr and has a sea level between 9 and 15 m above the present. In Barbados, Mesolella *et al.*[16] note a sea level 'above the present' at 125 kyr. In New Guinea, the Complex VIIb of Bloom *et al.*[17] has a sea level of +6 m at 125 kyr, and in the stable carbonate platform of Bermuda and in the Bahamas, Harmon *et al.*[18] and Neuman and Moore[19] recognize a +4 to +6 m level stand at 125 kyr. All these results are comparable to the Belle Hougue data in both timing and elevation of the last interglacial sea level maximum, thus allowing correlation of the 8-m raised beach of the western Channel with a similar sea level stand in the Mediterranean, Atlantic and Pacific. This correlation and radiometric date then suggests that the beaches of Andrews *et al.*'s[1] stage 2, and

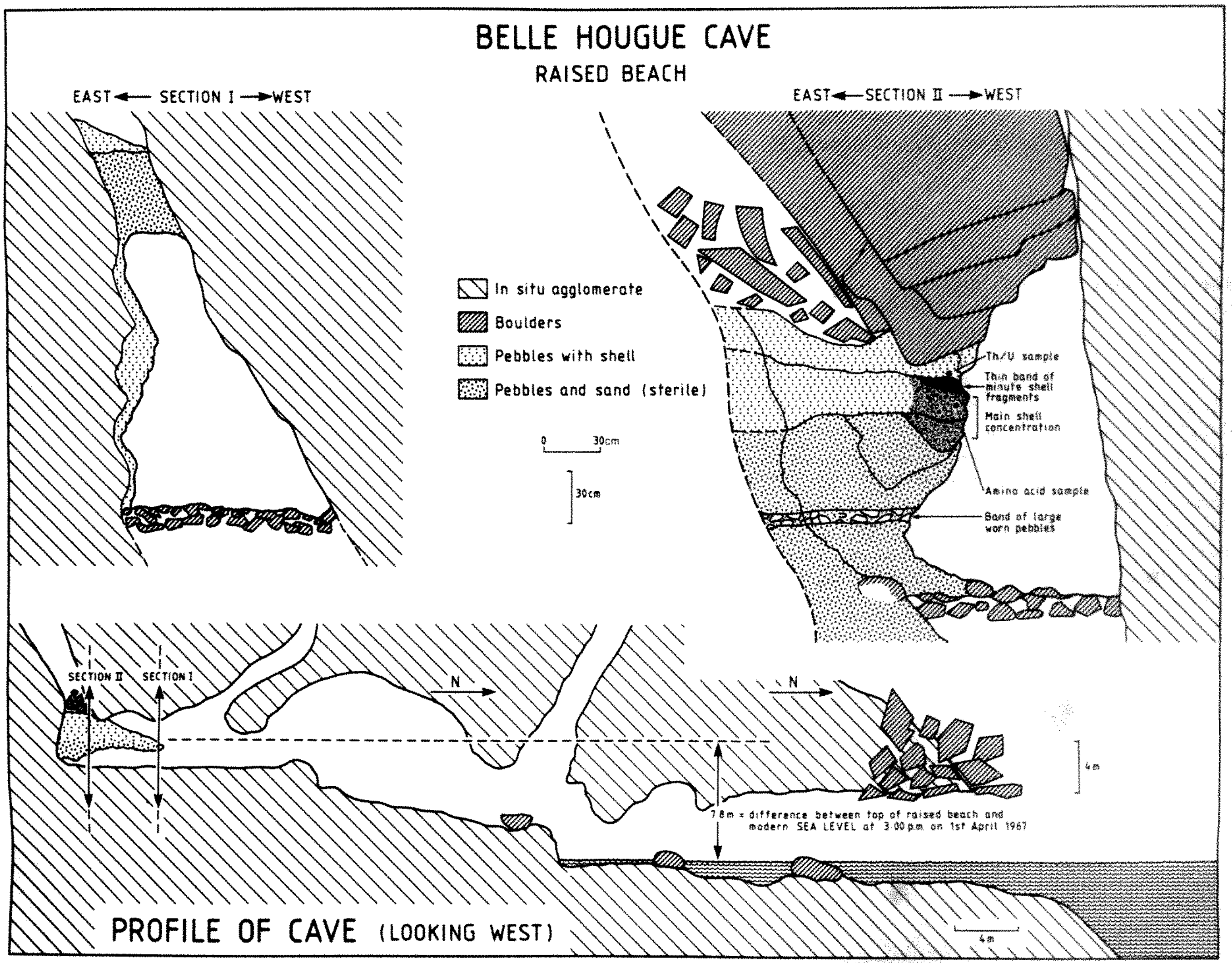

BELLE HOUGUE CAVE
RAISED BEACH
EAST ← SECTION I → WEST
EAST ← SECTION II → WEST
In situ agglomerate
Boulders
Pebbles with shell
Pebbles and sand (sterile)
0 30cm
30cm
Th/U sample
Thin band of minute shell fragments
Main shell concentration
Amino acid sample
Band of large worn pebbles
SECTION II
SECTION I
N
N
4m
7.8m = difference between top of raised beach and modern SEA LEVEL at 3.00 p.m. on 1st April 1967
PROFILE OF CAVE (LOOKING WEST)
4m

Fig. 2 Sections of the Belle Hougue Cave (based on a survey by J. Campbell, University of Cambridge, Department of Archaeology, 1967).

beaches occupying a similar altimetric range in the Channel Islands[3] may be pre-Ipswichian (stage 5) date.

We thank the University of Colorado Council for Research and Creative work for financing the amino acid values, Dr J. T. Renouf for assistance in securing the samples for analysis, Professor J. Campbell for permission to use the unpublished survey on which Fig. 2 is based, and Dr J. T. Hollin for comments. The Isotope Geology Unit of the SURRC is supported by a grant from the NERC.

Received 22 September; accepted 11 November 1980.

1. Andrews, J. T., Bowen, D. Q. & Kidson, C. *Nature* **281**, 556–558 (1979).
2. Miller, G. H., Hollin, J. T. & Andrews, J. T. *Nature* **281**, 539–543 (1979).
3. Keen, D. H. *Inst. geol. Sci. Rep.* 78/26 (1979).
4. Mottershead, D. N. *Geol. J. Spec. Iss.* No. 7, 299–321 (1977).
5. Elhai, H. *La Normandie occidentale entre la Seine et la Golfe normand-breton, étude morphologique*, Bordeaux (1963).
6. Zeuner, F. E. *A. Bull. Soc. Jersiaise* XIV, 238–254 (1946).
7. Harmon, R. S., Thompson, P., Schwarcz, H. P. & Ford, D. C. *Bull. natn speleol. Soc.* **37**, 21–33 (1975).
8. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39–55 (1973).
9. Phillips, L. *New Phytol.* **73**, 589–604 (1973).
10. Zeuner, F. E. *The Pleistocene Period* (Hutchinson, London, 1959).
11. Kidson, C. *Geol. J. Spec. Iss.* No. 7, 257–299 (1977).
12. Baden-Powell, D. F. W. *Proc. malacol. Soc.* **19**, 67–76 (1931).
13. Bowen, D. Q. *Proc. geol. Ass.* **84**, 249–272 (1973).
14. Stephens, N. in *The Glaciations of Wales and Adjoining Regions* (ed. Lewis, C.) (Longmans, London, 1970).
15. Butzer, K. W. in *After the Australopithecines* (eds Butzer, K.W. & Isaac, G. Ll. (Mouton, The Hague, 1975).
16. Mesolella, K. J.' Matthews, R. K., Broecker, W. S. & Thurber, D. L. *J. Geol.* **77**, 250–275 (1969).
17. Bloom, A. L. Broecker, W. S., Chappell, J. S., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 185–205 (1974).
18. Harmon, R. S., Schwarcz, H. P. & Ford, D. C. *Quat. Res.* **9**, 205–218 (1978).
19. Neuman, A. C. & Moore, W. S. *Quat. Res.* **5**, 215–227 (1975).

Relative departures in July temperatures in northern Canada for the past 6,000 yr

J. T. Andrews, P. T. Davis, W. N. Mode, H. Nichols & S. K. Short

Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado 80309

There has been concern about recent temperature trends and the future effects of CO_2 concentrations in the atmosphere[1,2]; but instrumental records only cover a few decades to a few centuries and it is essential that proxy data sources, such as pollen spectra from peats and lake sediments, be carefully interpreted as climate records. Several workers have shown statistically significant associations between the modern pollen rain and climatic parameters, an approach that by-passes the recognition of pollen/vegetation units. Statistically defined equations that associate abiotic and biotic elements are called transfer functions. We report here on the application of transfer function equations to nine middle and late Holocene peat and lake sediment sequences from northern Canada (Fig. 1).

Two transfer function methods are used[3]; the first follows the work of Imbrie and Kipp[4], whereas the second uses multiple stepwise regression to associate present climatic estimates with pollen percentages[5–9] directly. Modern pollen spectra from two transects are used; one trends SW–NE from south of the treeline in Keewatin to Clyde River, Baffin Island (ref. 9 and H. Nichols unpublished data) (KWA transect), whereas the other is oriented N–S from Fort Chimo in Labrador to Clyde River (ECA transect) (Fig. 1)[10,11]. Along the KWA transect, Nichols collected ⩾ 10 replicate mosses at 29 sites. The ECA transect consists of 87 surface samples from five major collecting localities (Fig. 1)[10,11].

Average July temperatures vary between 13 °C and 4 °C along the transects. Pollen data were regressed against this parameter. Pollen percentages were initially computed for all taxa. Many taxa were rare, hence percentages were recalculated on the basis of 13 common taxa for the Keewatin and for 19 taxa for the eastern Canadian Arctic transects. Thus, only common pollen types were used in the transfer function equations. Relevant data are listed in Table 1. We have compared the results from the Imbrie–Kipp and multiple regression approaches and shown that the two methods yield comparable results[6,8,9]. Equation MKWA-1B is based on the mean pollen percentages at the 29 sites, whereas MKWA-1 is calculated on 293 individual samples. All results indicate that mean July temperatures may be estimated on the basis of only a few taxa. Standard errors on the residuals are ⩽ ±1 °C (Table 1).

We applied our transfer function equations to a series of sites (1) to compare broad temperature trends; and (2) to recognize andcorrelate high frequency events. All sites that we use have been described previously[12–14]. Nichols[12] interpreted the diagrams from Ennadai, Thompson Landing and the two sites at Coppermine (Fig. 1) in terms of fluctuations in treeline and average July temperature. Transfer function equations MECA and MKWA were applied to lake/peat sites from Labrjor/Baffin Island and Keewatin, respectively. The resulting temperature estimates were smoothed. Finally, relative departures of July temperature were computed for each time-series as: $(x-\bar{x})/s$, where $\bar{x}$ and s are the mean and standard deviations (Table 2), respectively. Oscillations in temperature estimates could be simply 'noise'; thus, we examined residuals from time/temperature linear correlation models by the Runs Test, the null hypothesis was rejected in all cases but Thompson Landing. These results imply that the runs of temperature departures on Fig. 2 are not spurious.

Graphs of relative departures of average July temperature from their means are shown on Fig. 2. We only consider the past 6,000 yr so that changes in the pollen spectra primarily reflect climatic stresses rather than migrational histories of the different taxa. This statement may not be accurate for Labrador where an unusually late migration of conifers several thousand years after deglaciation took place[13–15].

The most important information in the time-series is the negative trend in the relative departures of July temperatures that begins between 3,000 and 2,500 yr BP.

Table 1 Information on the transfer function equations

	ECA-1	MECA-1B	MECA-2	KWA-1	MKWA-1B	MKWA-1
No. of samples	18	18	87	29	29	293
No. of taxa entered	19	19	19	13	13	13
No. of taxa selected in the equation	4	4	4	4	4	4
Multiple *r*	0.9	0.92	0.76	0.95	0.95	0.89
Standard error (±)	0.93	0.86	1.3	0.75	0.79	1.06
Names of the critical taxa	NA	*Alnus* *Picea* *Salix* Gramineae	*Alnus* *Picea* *Betula* Rosaceae	NA	*Alnus* *Picea* *Betula* Gramineae	*Alnus* *Picea* *Betula* Gramineae
Reference	6	6	6	9	9	9

The pollen taxa are in percentages. The climatic parameter estimated is July average temperature (°C). MECA and MKWA designate transfer function equations based on multiple, stepwise regression analysis. The Imbrie–Kipp approach has also been used and gives comparable results[2,8]; these results are designated ECA and KWA. NA, not applicable.

0028-0836/81/030164—04$01.00

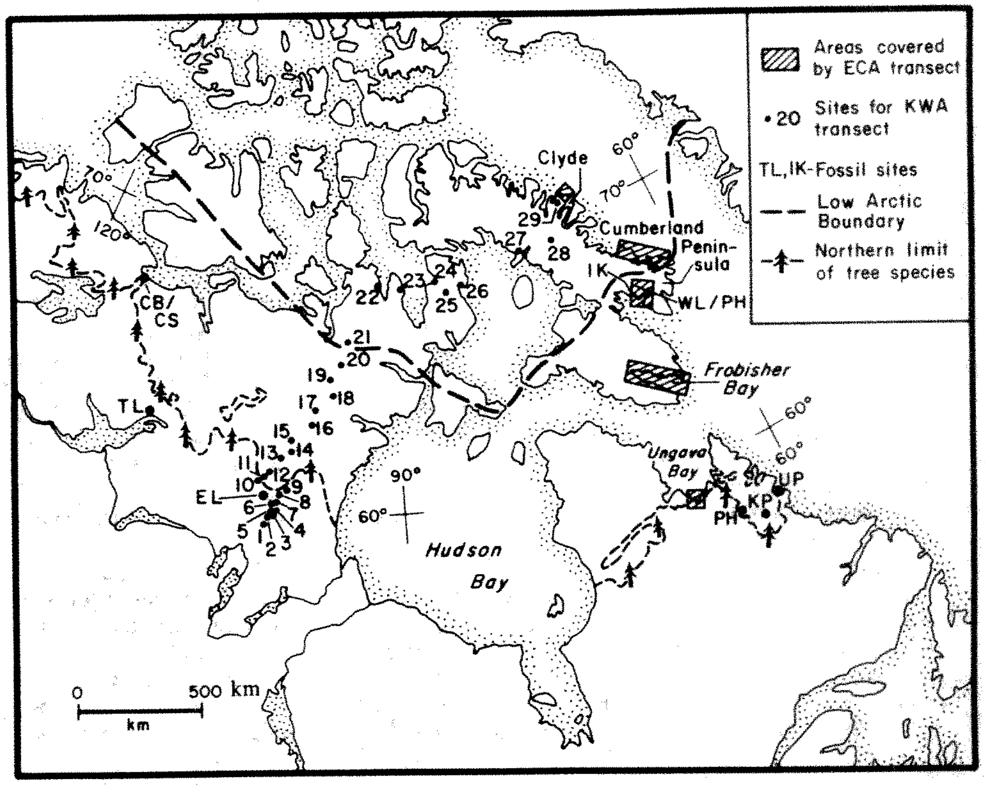


Fig. 1 Location of the two transects of modern surface pollen samples. Keewatin to Baffin Island sites numbered 1–29; Labrador to Baffin Island sites outlined as rectangles. Treeline and the Low Arctic vegetation boundary shown. Peat sites: EL, Ennadai Lake; TL, Thompson Landing; CB and CS, Coppermine Beach and Saddle; WL/PH, Windy Lake/Pass Head. Lake sediment sites: IK, Iglutalik Lake; UP, Ublik Pond; KP, Kogaluk Plateau Lake; PH, Pyramid Hills lake.

On Baffin Island, both records show a statistically significant decrease in average July temperature over the period of record (Fig. 2). The local thermal maximum terminated between 5,300 and 5,000 yr ago. The lowest estimated July departure during the past few thousand years occurred between 600 and 200 yr ago, although both records indicate continuously cool summers after ~2,300 yr BP. Present July temperatures are closely predicted by modern surface pollen spectra for this region; hence, the present temperature estimates are as warm, if not warmer, than any values estimated for the past 1,000–2,000 yr. The ages for the Windy Lake peat are based on a uniform sedimentation model. Davis[11] argued that a nonlinear relationship better fitted the 14 ^{14}C dates. Figure 2 shows alternative positions for peaks and troughs of temperature departures. However, the uniform model gives a satisfactory agreement with Iglutalik Lake.

In Labrador[13], the Ublik Pond site is our oldest record, and the July temperature departures >6,000 yr BP might be influenced by migrational lags in vegetation. The transfer functions predict that the major late Holocene cooling did not begin until 2,500 yr ago, and that the local thermal maximum occurred 4,000–3,000 yr BP. The other two sites from Labrador support this interpretation[13,14]. However, the record from Kogaluk Plateau Lake ends at ~2,500 yr BP, so that late Holocene changes are not recorded at that site.

Four sites have been examined from Keewatin[12]. Data from the Coppermine Beach showed no statistically significant trend of departures against age. The Beach site is on a dry, raised beach, whereas the Coppermine Saddle site is wetter.

Most of our sites (Fig. 2) indicated that July temperature departures became, on average, more negative over the last few thousand years. However, this does not preclude rapid climatic changes at certain times. Nichols[12,16] argued that the pollen evidence from Keewatin favoured a series of synchronous step changes in climate across the mainland of arctic Canada, which were superimposed on a general decline in July temperatures. Certain temperature graphs from our work show evidence for similar rapdi changes (Fig. 2). In particular, the very rapid

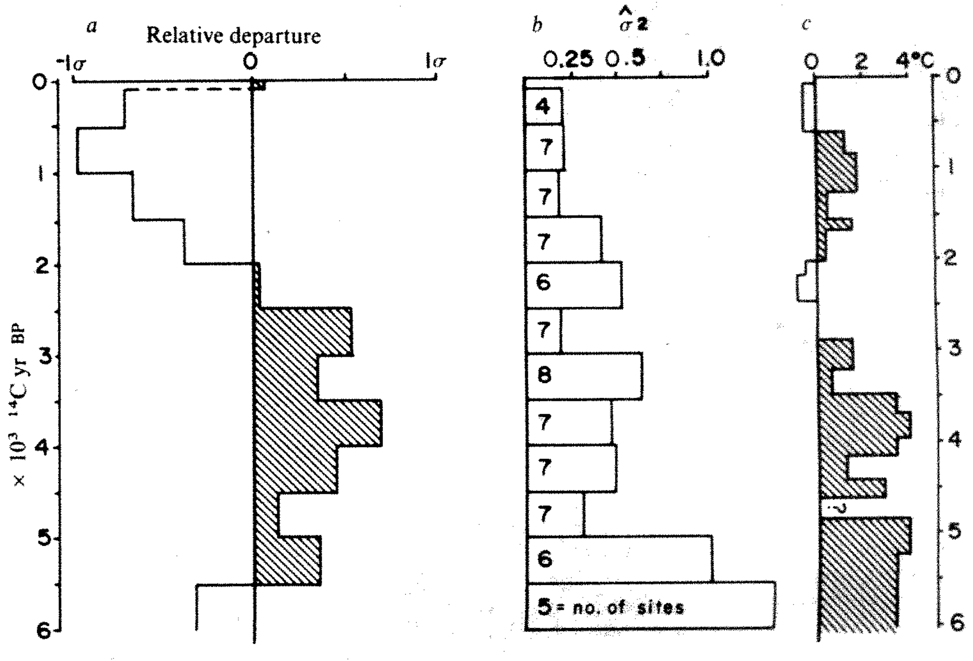


Fig. 2 Time-series of estimated July temperatures at nine sites across northern Canada. Temperatures expressed as relative departures from the smoothed mean of each series. Table 2 lists $\bar{x}$ and s for each series. A–A^1, and so on, refer to different timings of peaks at Windy Lake depending on which sedimentation model is used.

SCIENCE COLLEGE
Upper Circular Road
CALCUTTA

Table 2 Analysis of temperature data in Fig. 1

Site	Ublik Pond	Kogalu Pl.	Pyramid L.	Thompson L.	Ennadai L.	Coppermine Beach	Coppermine Saddle	Windy L.	Iglutalik
$\bar{x}$	12.1	13.1	13.2	10.0	9.4	9.25	10.1	6.3	8.0
s	0.7	0.7	0.5	0.8	1.2	0.7	1.2	1.0	0.8
n	22	14	34	59	40	20	23	90	53
Present July temperature (°C) (±)	11±	13	13	12.8	12.3	9.3	9.3	7	7

$\bar{x}$ and s are mean and standard deviation for each time series.

cooling that began in the late Holocene is notable. At Coppermine Saddle, a temperature drop of about 2 °C in 150 yr between 1,400 and 1,250 yr BP is estimated. At Ennadai Lake, Windy Lake and Ublik Pond, July temperatures fell 1–2 °C close to 2,500 yr BP.

The timing and direction of these changes were determined by averaging the data (Fig. 2) in 500-yr time-blocks and expressing the departures as ± values about the mean of each series (Fig. 3). If regionally offsetting departures were present, we would expect the average departure to be close to zero. However, regionally synchronous changes in estimated July temperatures can be identified on Figs 2 and 3. Our study area extends across the region covered by one of the main upper atmospheric troughs. As this trough shifts to the east of west, out-of-phase climatic trends may occur, at least on short time scales[17,18]. However, arguments for and against synchrony of climatic change on the longer time scales of centuries occur[12,18,19].

Between 5,500 and 6,000 yr ago, in northern Canada, mean July temperature departures were, on average, negative. July temperature departures then rose to above average values until 2,500 yr BP, with a reversal between 5,000 and 4,500 yr BP (Fig. 3). The highest average departures from July conditions occurred 3,500–4,000 yr BP, which apparently represents a Holocene local thermal optimum for the eastern and central Canadian Arctic. Between 3,000–2,500 and 1,000–500 yr BP, the average departures declined steadily and reached a trough in the latter interval.

Nichols'[12] graph of departures of July temperatures from modern values is reproduced on Fig. 3. The major difference between that reconstruction and ours is the interpretation of events after 2,500 yr BP; our reconstruction shows a steady decline to values well below the mean of the nine time-series, whereas Nichols' reconstruction shows less of a decline and, in fact, has temperatures above present between 2,000 and

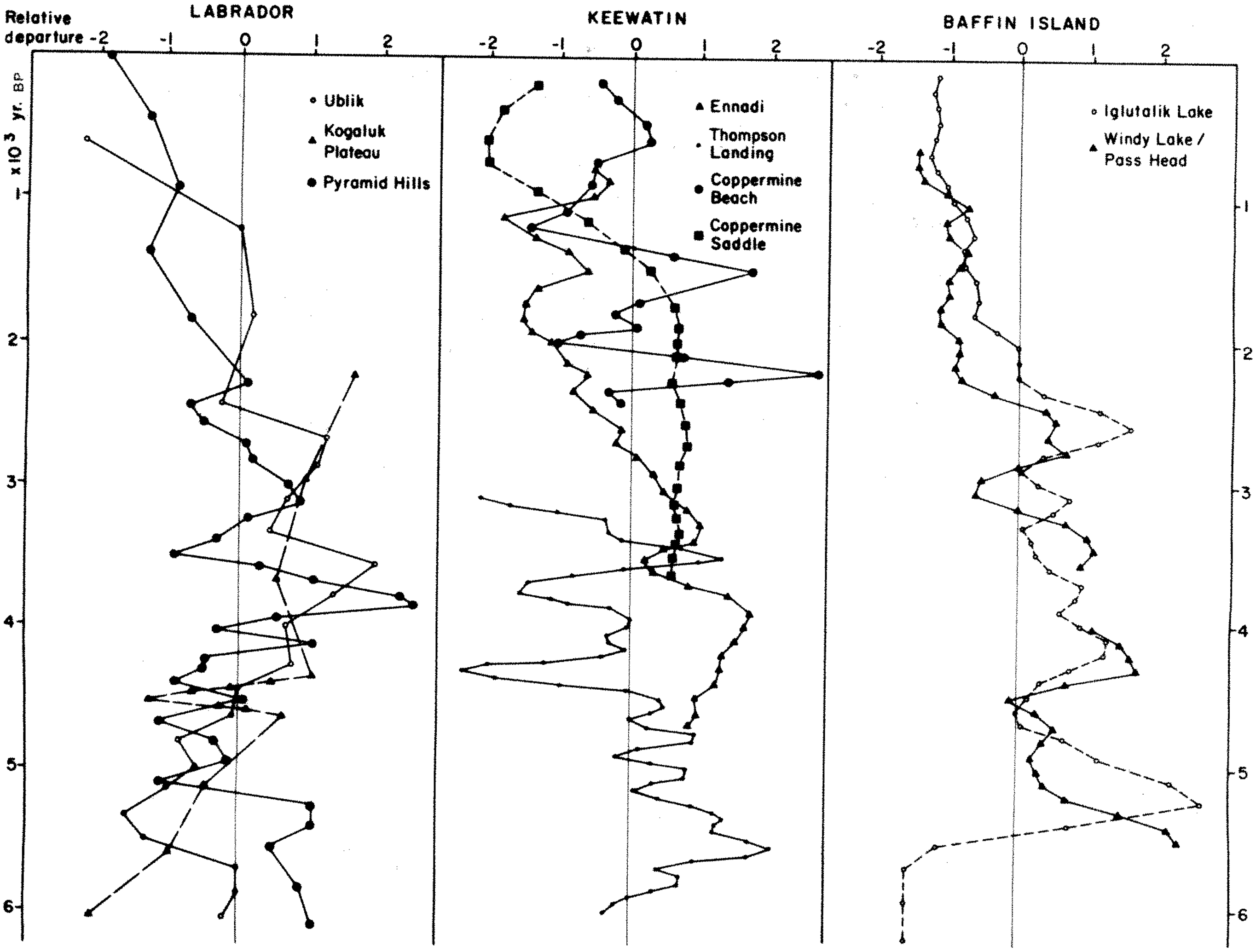


Fig. 3 *a*, Average relative departures for all sites for 500-yr time blocks; shaded area indicates period when relative departures were above the average July temperature (°C) for the past 6,000 yr. *b*, Variance of the relative departures, number within each block represents the number of curves available for each interval. *c*, Nichols' model of variations in July temperature (°C) for central, northern Canada based on an analysis of fossil pollen spectra. The temperature history of the past 200 yr is inferred and based on comparison of present average July temperatures compared to the fossil record.

600 yr BP. We have yet to examine our data for how far the variance of Fig. 3*c* is due to pronounced regional differences in our estimated July temperature records, or how far it is associated with uncertainties in the pollen transfer functions, or how much is associated with problems of ^{14}C dating organic poor sediment[13,20].

Our data indicate that the most useful northern analogue for a 'warm' summer world during the past 6,000 yr is between 3,500 and 4,000 yr ago. Figures 2 and 3 indicate that July temperatures in the central and eastern Canadian Arctic departed, on average, by about 0.75 standard deviation from the 6,000-yr mean, or ~1 °C. Locally, estimated July temperature reconstructions are predicted to have been 2 °C to 3 °C warmer than at present.

This research was supported by NSF grant ATM-77-17459. We thank Dr L. D. Williams for helpful comments.

Received 14 July; accepted 8 November 1980.

1. Wigley, T. M. L., Jones, P. D. & Kelly, P. M. *Nature* **283,** 17–21 (1980).
2. Manabe, S. & Stouffer, R. J. *Nature* **282,** 491–493 (1979).
3. Webb, T. III & Bryson, R. A., *Quat. Res.* **2,** 70–115 (1972).
4. Imbrie, J. & Kipp, N. G. in *Late Cenozoic Ice Ages* (ed. Turkeian, K. K.) 71–181. Yale University Press, 1971).
5. Kay, P. A. *Quat. Res.* **11,** 125–140 (1979).
6. Andrews, J. T., Mode, W. N. & Davis, P. T. *Arc. Alp. Res.* **12,** 41–64 (1980).
7. Howe, S. & Webb, T. III *5th Conf. Probability and Statistics* 152–157 (American Metallurgical Society, 1977).
8. Andrews, J. T. *et al. Geol. Soc. Am. Abstr.* **11,** 379–380 (1979).
9. Andrews, J. T. & Nichols, H. (in preparation).
10. Mode, W. N. thesis, Univ. Colorado, Boulder (1980).
11. Davis, P. T. thesis, Univ. Colorado, Boulder (1980).
12. Nichols, H. *INSTAAR Occ. Pap. No. 15* (University of Colorado, 1975).
13. Short, S. K. thesis, Univ. Colorado (1978).
14. Short, S. K. & Nichols, H. *Arc. Alp. Res.* **9,** 265–290 (1977).
15. McAndrews, J. H. & Samson, G. *Geogr. Phys. Quat.* **31,** 177–184 (1977).
16. Nichols, H. *Eiszeitalter. Genenw.* **18,** 155–167 (1967).
17. Barry, R. G. *et al. Arc. Alp. Res.* **9,** 193–210 (1977).
18. Keen, R. A. *INSTAAR Occ. Pap. No. 34* (University of Colorado, 1980).
19. Short, S. K. & Andrews, J. T. *Geogr. Phys. Quat.* (in the press).
20. Stuckenrath, R., Miller, G. H. & Andrews, J. T. *Arc. Alp. Res.* **11,** 109–120 (1979).

Footprints of a Pleistocene hominid in northern Kenya

Anna K. Behrensmeyer

Peabody Museum and Department of Anthropology,
Yale University, New Haven, Connecticut 06520

Léo F. Laporte

Earth Sciences Board, University of California, Santa Cruz,
California 95064

We report here that in 1978, while studying Koobi Fora Formation sedimentary environments and fossils in northern Kenya, we discovered single footprints and trails of large vertebrates in a number of different strata. One of these trails is that of a bipedal hominid, the second known record of early hominid tracks in Africa, the other being that reported at Laetoli, Tanzania[1].

The Koobi Fora Formation is exposed for 2,600 km^2 along the northeastern shore of Lake Turkana. It is up to 325 m thick and consists of laterally variable facies deposited in lacustrine, lake margin and fluvial environments[2,3]. Within the Upper Member of the Koobi Fora Formation, in the vicinity of Koobi Fora (Fig. 1*a*), is a poorly consolidated 10–20-m sedimentary complex of tuffaceous sands and silts designated the Koobi Fora Tuff[4]. This has yielded a scatter of radiometric dates with a preferred estimate of 1.5–1.6 Myr (ref. 5). Correlation based on lithology and stratigraphic position can be made from the Koobi Fora Tuff to the Okote Tuff (Karari area) and Lower/Middle Tuff (Ileret area) to the north[2,3], although outcrops between areas are discontinuous (Fig. 1*a*). The Okote and Lower/Middle Tuffs are also dated at 1.5–1.6 Myr (ref. 5). Biostratigraphic correlations to other dated sequences including the Omo Shungura Formation[6] and Olduvai Gorge[7,8] support this date.

Table 1 Numerical data for the Koobi Fora Tuff hominid tracks

Hominid print	Width (cm)	Length (cm)	Maximum depth (cm)	Orientation	Notes
L1	13.0	26	6.0	80°	Deepest near first toe, morphology poorly preserved
R1	18.0	32	6.0	90°	Deepest under ball of foot, morphology poorly preserved
(L2)	—	—	—	—	Missing, probably due to locally hard substrate
R2	7.0	—	1.5	~87°	Incomplete print, probably heel only
L3	10.5	27	4.0	51°	Includes partial toe prints, deepest at heel
R3	10.0	26	3.0	92°	Probable minor slipping during step
L4	—	—	—	~88°	Partial, superimposed on composite hippopotamus prints
R4	9.5	25	2.5	92°	Greatest depth at heel and along outside edge

R = right foot, L = left foot; numbered from west to east in the direction of movement. Width is measured at the ball of the foot, length from toe to heel along the midline.

There are extensive exposures of the Koobi Fora Tuff in Area 103 (lat 3° 53′N, long 36° 15′E), 13 km SSE of the Koobi Fora research camp, in a series of fault blocks. In 1978, step trenches revealed over 20 distinct strata with sedimentary structures which, in cross-section, seemed to be imprints of large vertebrate feet on originally soft, muddy substrates[9]. The sediments record marginal lacustrine environments and are vertically complex, with alternating thin units (1–15 cm) of sand and mud. Within the uppermost 1 m, cross-sections of particularly well preserved footprints were found in a unit of sandy mud (Fig. 1*b*). An area of 1 m^2 of the upper surface of this unit was cleared of overlying sediment to determine the shapes, in plan view, of the footprints. Excavation revealed several deep, four-toed impressions of hippopotamus feet and a shallow depression resembling a human footprint. Kimolo Mulwa and Bernard Kanunga of the National Museums of Kenya were responsible for the uncovering and recognition of this first hominid track (R4). Excavation of a second 1-m^2 area exposed two further complete human-like prints in line with the first, confirming the discovery of a hominid trail. Excavation in 1979 expanded the area of exposed bedding surface to 12 m^2 and revealed two further complete hominid tracks together with those of many other vertebrates (Fig. 1*c*).

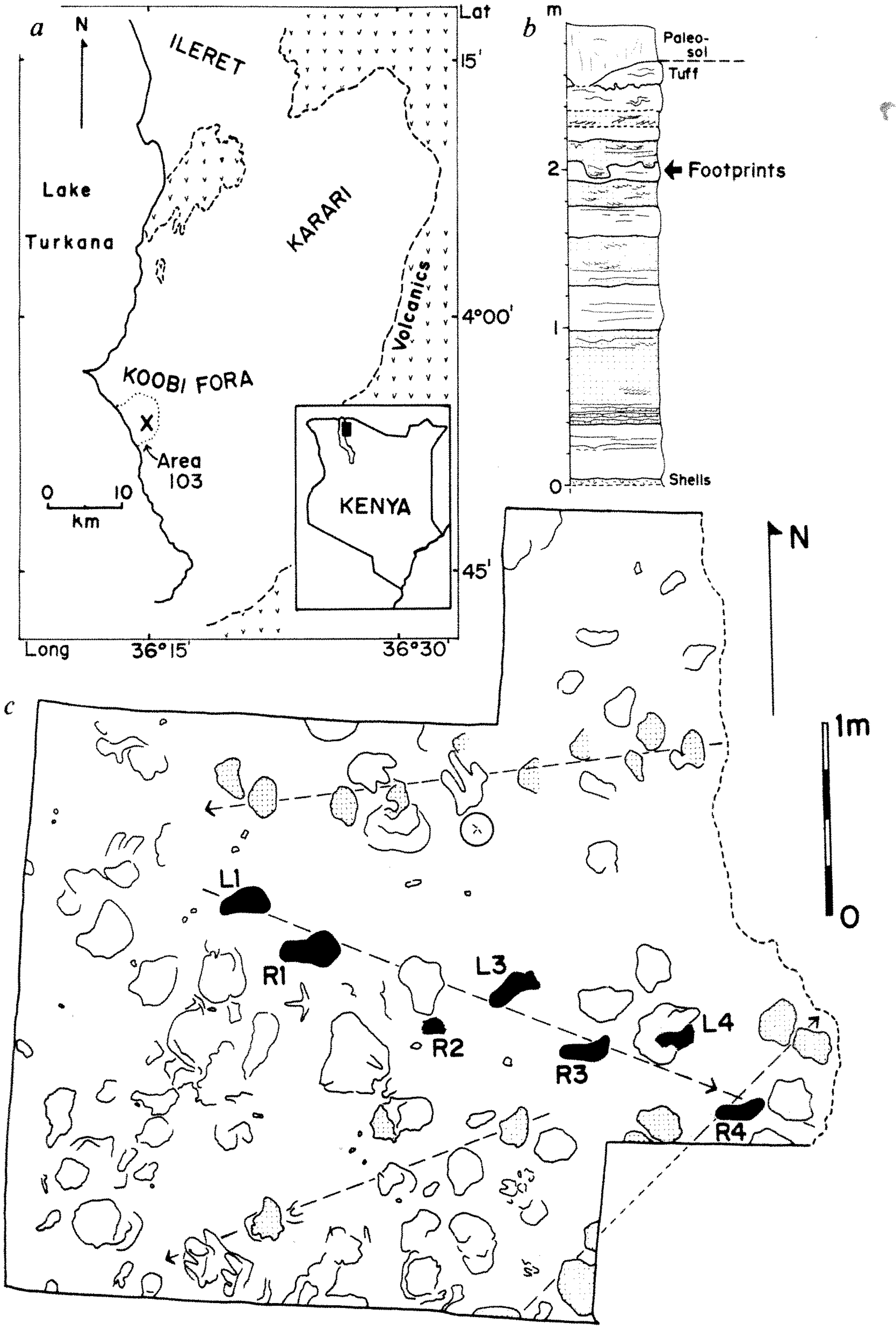
The footprints are in a homogeneous, pinkish-buff sandy mud (sand : silt : clay = 12 : 59 : 29) about 12 cm thick, interbedded with ripple-laminated, fine pumaceous sands of comparable thickness, with sharp upper and lower contacts. The underlying muddy sand unit has slight, concave-upward deformation of laminae beneath the hippopotamous footprints, which penetrate through the sandy mud to within 1 cm of its base. This deformation shows the effects of the weight of the animals and indicates that the sand under the sandy mud was relatively compact when trampled. Muddy sand (sand : silt : clay = 81 : 10 : 9) infillings of the footprints are slightly coarser than the overlying sand and include fine silt laminae, concentrations of heavy minerals, and in some cases abundant sand-coated mudclasts. The footprint layer has an irregular upper surface with higher rims around the deeper tracks, apparently a result of displacement of the mud as the foot penetrated the substrate. There seems to have been no overall scouring before burial of the mud; the sand-bearing currents simply filled and buried the tracks. The sandy mud was viscous enough to form thin, 10-cm-high walls between tracks, and it is doubtful that these would have survived even slight erosion by current action. There are no mudcracks in the footprint unit as evidence for subaerial exposure, although a sample of the sandy mud readily developed cracks when wetted and dried. We conclude that the footprints were formed when the mud surface was submerged under quiet

0028–0836/81/030167—03$01.00

water. The track of a small wading bird indicates that the water was shallow, < 10 cm deep.

On the 12 m^2 of excavated surface, 89 distinct depressions between 2 and 23 cm deep can be identified as the tracks of large vertebrates (Fig. 1*c*). Of these, 22 indicate a foot similar to that of modern hippopotamus, and in 15 there are clear impressions of toes. In some cases the toe-marks show that the animals slipped while stepping in the mud. Three distinct trails record one hippopotamus going north-east and two going west. The larger prints (25–32 cm wide) can be attributed to the large fossil hippopotamus, *Hippopotamus gorgops*, recognized from the Koobi Fora Formation at that time[10]. The smaller ones (18–20 cm wide) could be due to juveniles of this species or may belong to the fossil pigmy species, *Hippopotamus aethiopicus*[10]. Four of the other, smaller tracks resemble those of modern large Bovidae (the size of a topi, *Damaliscus korrigum*, about 120 kg) but are not sufficiently well preserved for definite identification. There are also two distinct bird prints, one large and one small. The larger is a clear, three-toed impression 20 cm wide, similar in size and morphology to that of a modern bustard or crane. The smaller, which is 2.0 cm wide, resembles tracks of a small shore bird such as a plover.

Five of the depressions on the surface are complete tracks made by the feet of a bipedal hominid (Fig. 1*c*). Two additional



[illegible] SCIENCE COLLEGE [illegible] Upper Circular Road CALCUTTA

Fig. 1 Map showing location of hominid footprint locality (×) in relation to Lake Turkana and areas of outcrop of the Koobi Fora Tuff and its lateral equivalents in the north. Inset shows location in Kenya. *b*, Stratigraphic section showing upper 2.6 m of the Koobi Fora Tuff and the position of the footprint unit (arrow). Muddy sand units stippled, sandy mud blank. Heavy dashed line indicates top of Koobi Fora Tuff. *c*, Map of the excavated surface of the footprint layer, with hominid footprints (shown in black). Dashed lines indicate direction of movement for the more distinct trails (light stipple in hippopotamus footprints forming the trails). Track of small wading bird is circled.

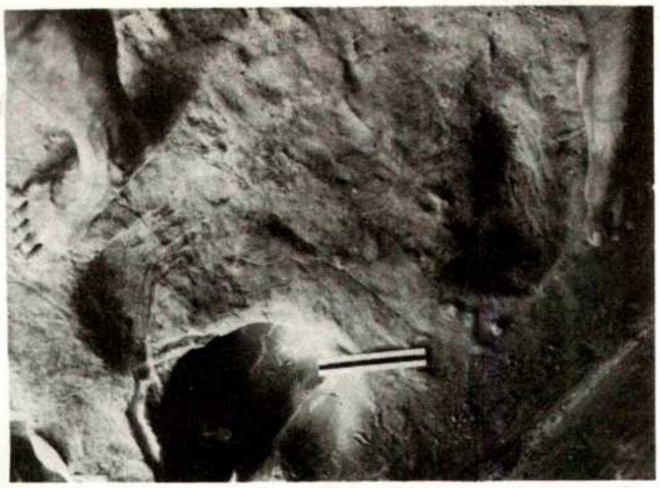


Fig. 2 Two of the best-preserved footprints in the 1.5-Myr old trail discovered in northern Kenya, framed by the modern feet of *Homo sapiens* (27 cm long). These prints are referred to as L3 and R3 in Fig. 1*c*. Scale bar, 15 cm. (Photograph by Léo F. Laporte.)

tracks can be attributed to the same individual although they would not be identifiable as hominid if observed separately from the trackway defined by the other five footprints. The complete tracks are shallower towards the east, in the direction of movement (Table 1). The westernmost tracks (L1 and R1) are oversized and poorly shaped compared with the others, and bear evidence of minor erosion and slumping before burial. The absence of an L2 track is puzzling, but may be due to a patch of substrate that was compact enough to bear the weight of the hominid or to local scouring before burial. More hominid tracks would probably be uncovered if the excavation were extended westward along the line of the trackway, although the increasing amount of overburden would make that a difficult operation.

From the dimensions of the best-preserved tracks, L3, R3 and R4, we estimate that the hominid foot was about 26 cm long and 10 cm wide (Table 1). These fall close to average dimensions for American *Homo sapiens* males[11]. Other tracks in the series are enlarged or incomplete and cannot be used to reconstruct foot shape. The height of the hominid can be estimated at 1.6–1.7 m based on the relationship of foot length to stature for American negro and white males[11], and 1.8 m based on those for San bushmen[12]. The hominid prints are longer and relatively narrower than those recorded at Laetoli, which average 18.5 cm long and 8.8 cm wide (trail 1) and 21.5 × 10 cm (trail 2)[1].

In the study of footprints, a standard definition of 'stride' is the distance from a point on one foot (left or right) to the same point on the next impression of the same foot[13]. The length of the stride of the hominid is about 80 cm, less than that of modern humans with feet of comparable size during normal walking[14]. This, along with the orientation of the feet, seems to indicate a hesitant, somewhat sideways progression across a slippery surface with one mis-step (L4) into a deep hippopotamus track, which may have been concealed by turbid water.

At least two bipedal hominid taxa are represented by fossil bones from the Koobi Fora Formation at the time of deposition of the footprint unit: *Homo erectus* and *Australopithecus robustus*[15–17]. In Area 103, only the remains of *H. erectus* have been found in the Koobi Fora Tuff although in the Karari and Ileret regions both taxa occur in units correlated with the Koobi Fora Tuff. The *H. erectus* specimens from the Koobi Fora Tuff occur within its lower 6 m; none has been found within the upper part where the hominid trackway is recorded. It is not possible to determine which hominid made the footprints.

Trails of early hominids known from the Pliocene Laetolil Beds in northern Tanzania were preserved in subaerial volcanic ash strata dated at 3.60–3.75 Myr (refs 1, 12). They occur with a large series of tracks and trails representing a wide diversity of savannah animals. The hominid tracks are thought to belong to the same species as hominid fossils from the Laetolil Beds, designated as *Australopithecus afarensis*[1,12,18]. In contrast, the Koobi Fora footprints occur on a muddy lake margin with tracks of a limited number of semi-aquatic vertebrates. The Koobi Fora footprints contribute a reference point on hominid foot morphology, locomotion, behaviour and ecology 2 Myr younger than the Laetoli occurrences and ~1 Myr older than late Pleistocene human footprints in Europe[19].

A latex mould and plaster cast of the Koobi Fora hominid footprints described here is stored at the National Museums of Kenya, Nairobi. We thank the Government of Kenya, the National Museums of Kenya and the International Louis Leakey Memorial Institute for African Prehistory for their cooperation and support. Richard Leakey, Meave Leakey, Glynn Isaac, Kamoya Kimeu and Tim White are acknowledged for their suggestions and help. The excavation crew, headed by Muteti Nume and Mukilya Mun'goka, were directly responsible for the technical work that led to the discovery of the trackways. Hilde Schwartz and Mahmood Raza assisted with the excavation work. Palaeoecological research was sponsored by NSF grant EAR77-23149 and is a part of the overall Koobi Fora Research Project, which is supported by grants from the National Geographic Society, the NSF and the W. H. Donner Foundation.

Received 7 July, accepted 4 November 1980.

1. Leakey, M. D. & Hay, R. L. *Nature* **278,** 317–323 (1979).
2. Findlater, I. C. in *Koobi Fora Research Project* Vol. 1 (eds. Leakey, M. G. & Leakey, R. E.) 14–31 (Clarendon, Oxford, 1978).
3. Vondra, C. F. & Bowen, B. E. in *Geological Background to Fossil Man* (ed. Bishop W. W.) 395–414 (Scottish Academic, Edinburgh, 1978).
4. Findlater, I. C. in *Earliest Man and Environments in the Lake Rudolph Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. Ll. & Leakey, R. E. 94–104 (University of Chicago Press, 1976).
5. Fitch, F. J. & Miller, J. A. in *Earliest Man and Environments in the Lake Rudolph Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. Ll. & Leakey, R. E.) 123–147 (University of Chicago Press, 1976).
6. Brown, F. H. Howell, F. C. & Eck, G. G. in *Geogological Background to Fossil Man* (ed. Bishop W. W.) 473–498 (Scottish Academic, Edinburgh, 1978).
7. Cerling, T. E., Brown, F. H., Cerling, B. W., Curtis, G. H. & Drake, R. E. *Nature* **279,** 118–121 (1979).
8. Harris, J. M. & White, T. D. *Trans Am. phil. Soc.* **69,** 1–128 (1979).
9. Laporte, L. F. & Behrensmeyer, A. K. *J. Sedim. Petrol.* **50** (in the press, 1980).
10. Harris, J. M. in *Koobi Fora Research Project* Vol. 1 (ed. Leakey, M. G. & Leakey, R. E.) 32–63 (Clarendon, Oxford, 1978).
11. Hrdlička, A. *Am. J. phys. Anthrop.* **12,** 15–33 (1928).
12. White, T. *Science* **203,** 175–176 (1980).
13. Sargeant, W. A. in *The Study of Trace Fossils* (ed. Frey, F. W.) 283–324 (Springer, New York, 1975).
14. Murray, N., Drought, A. B. & Kory, R. C., *J. Bone Jt Surg.* 46-A (2), 335–360 (1964).
15. Walker, A. & Leakey, R. E. *Scient. Am.* **29**(8), 54–66 (1978).
16. Leakey, R. E. *Nature* **231,** 241–245 (1971).
17. Day, M. H. in *Early Hominids of Africa* (ed. Jolly, C.) 311–346 (Duckworth, London, 1978).
18. Johanson, D. C. & White, T. D. *Science* **203,** 321–350 (1979).
19. Oakley, K. P., Cambell, B. G. & Molleson, T. I. (eds) *Catalogue of Fossil Hominids Part II: Europe*, 349–350 (British Museum (Nat. Hist.) publ. 711, 1971).

Relationship between fungus and alga in the lichen *Cladonia cristatella* Tuck

Vernon Ahmadjian* & Jerome B. Jacobs†

*Department of Biology, Clark University, Worcester,
Massachusetts 01610
†St Vincent Hospital and University of Massachusetts Medical
School, Worcester, Massachusetts 01605

The nature of the lichen symbiosis is not clear. It is generally thought to be mutualistic but this concept is not supported by experimental evidence[1]. Early workers[2,3] considered that lichens represented algae parasitized by fungi—as evidence, they noted algal cells in a lichen thallus that were dead or penetrated by fungal haustoria. Others, however, cited the seemingly healthy and long-lasting nature of lichens as evidence of mutualism. As we report here, our observations of artificial syntheses of the mycobiont *Cladonia cristatella* ('British soldiers') with different algae suggest that the relationship in this lichen is one of controlled parasitism. The mycobiont formed squamules mostly with algae related to its natural phycobiont, an indication perhaps of a long period of co-evolution between the symbionts of this lichen.

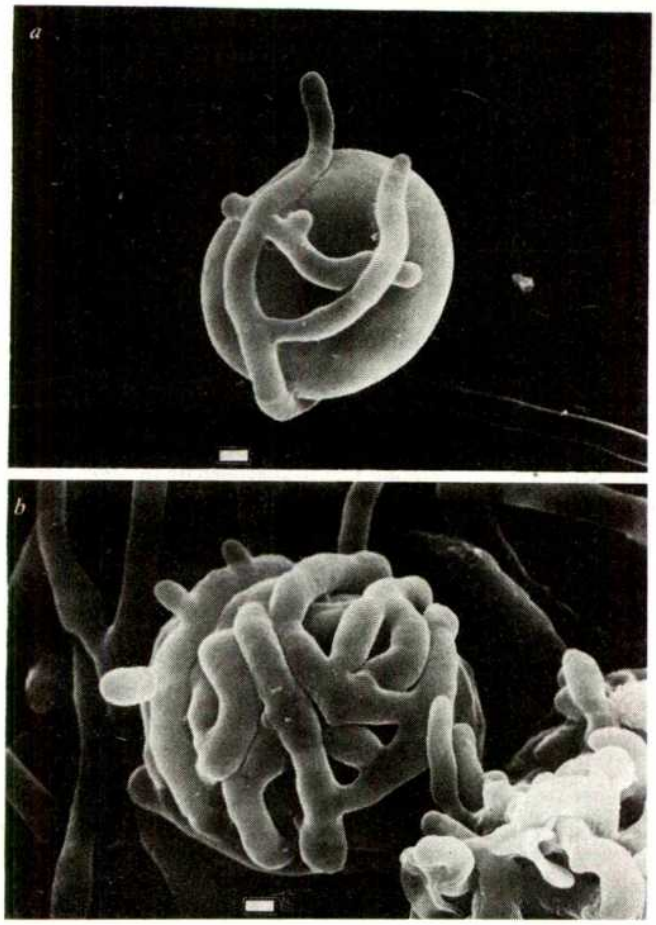


Fig. 1 Envelopment of cells of *Trebouxia erici* by mycobiont. With the aid of a dissecting microscope, mica strips with colonies of resynthesized lichens were cut into 13-mm pieces and placed in Petri dishes filled with 2% glutaraldehyde in 0.1 M phosphate buffer, pH 7.2, and fixed for 30 min. The colonies were rinsed briefly in phosphate buffer and then fixed for 30 min at room temperature in 1% OsO_4 in 0.1 M phosphate buffer. The specimens were rapidly dehydrated through an ascending alcohol series and dried by the critical point method. Specimens were cemented to specimen stubs, sputter coated with approximately 100 Å of gold, and examined at 20 KV in a JEOL JSM 35U scanning electron microscope. *a*, Algal cell encircled by branches from one hyphal filament. A thin gelatinous matrix around the alga binds the hyphae. *b*, Algal cell almost fully enveloped by hyphae. Haustorium has penetrated the cell at the top and along one hypha are three new branches. Scale bars, 1 μm.

Procedures used are described in Table 1 legend. After 2–3 weeks of axenic culture, the fungus formed squamules with five species of *Trebouxia* phycobionts and presquamules with *Trebouxia italiana* and *Friedmannia israeliensis*. The phycobionts were isolated from 13 different lichens. The fungus did not form lichenized structures with 17 algal symbionts, the free-living algae *Chlorella vulgaris* and *Pleurastrum terrestre* or unlichenized representatives of the Chlorococcales (Table 1). The fungus was not specific in its attachment, encircling glass beads (10–15 μm in diameter) in the same manner as algal cells, although the hyphae could be easily dislodged from the beads.

Algal cells were covered by a substance which bound the fungal hyphae (Fig. 1*a*) similar to that seen between the symbionts of the natural lichen and in a different lichen synthesized previously[4]. When initial contacts were established, the fungus, depending on the alga contacted, formed lichenized associations or destroyed the cells with its haustoria. Initial contacts (Fig. 1*a*) were followed by further encirclement of the algae, some cells being almost completely enveloped (Fig. 1*b*). After the fungus had secured small groups of algal cells and formed presquamules (Fig. 3*a*), a network of hyphae formed over the presquamules (Fig. 3*b*) linking them into a common structure. Thickened and fused hyphae with irregular extra-hyphal material developed first in small patches and later over a broader area of the mixed culture, forming the upper cortex and network of hyphae which covered the squamules. Further differentiation resulted in an algal layer and medulla. The morphology and dimensions of the thallus layers were identical to those of the natural squamules.

Table 1 Interactions between the mycobiont *Cladonia cristatella* (s.s.31) and various algae

Algae lichenized by mycobiont
*Friedmannia israeliensis**
(UTEX 1181; non-lichenized,
presquamules only)
Trebouxia erici
(UTEX 910; *Cladonia*
cristatella)
Trebouxia excentrica
(*Cladonia bacillaris*
Cladonia leporina
Cladonia subtenuis
Lecidea metzleri
Lepraria sp.
UTEX 1714; *Stereocaulon*
dactylophyllum
var. *occidentale*)
Trebouxia glomerata
(*Cladonia boryi*
Huilia albocaerulescens
UTEX 894; *Stereocaulon*
saxatile)
Trebouxia italiana
(CCAP 219/5b; *Xanthoria*
parietina,
presquamules only
Trebouxia magna
(UTEX 902; *Pilophoron*
acicularis)
Trebouxia pyriformis
(*Stereocaulon pileatum*)

Algae parasitized by mycobiont
*Bracteacoccus minor**
(UTEX 66; non-lichenized)
Chlorella vulgaris
(non-lichenized)
*Chlorococcum ellipsoideum**
(UTEX 972; non-lichenized)
*Chlorococcum gelatinosum**
(UTEX 1773; non-lichenized)
Myrmecia biatorellae
(*Dermatocarpon tuckermanii*)
*Nautococcus pyriformis**
(UTEX 125; non-lichenized)
*Nautococcus terrestris**
(UTEX 1794; non-lichenized)
*Neochloris terrestris**
(UTEX 947; non-lichenized)
Nostoc punctiforme
(*Peltigera canina*)
Pleurastrum terrestre
(UTEX 333; non-lichenized)
Pseudochlorella sp.
(*Lecidea granulosa*
Lecidea scalaris
Stereocaulon strictum)
Pseudotrebouxia aggregata
(*Lecidea fuscoatra*
UTEX 180; *Xanthoria* sp.)
Pseudotrebouxia corticola
(UTEX 909; non-lichenized)
Pseudotrebouxia decolorans
(UTEX 781; *Buellia punctata*)
Pseudotrebouxia galapagensis
(*Ramalina* sp.)
Pseudotrebouxia gigantea
(*Caloplaca cerina*)
Pseudotrebouxia higginsiae
(*Buellia straminea*)
Pseudotrebouxia impressa
(UTEX 892; *Physcia stellaris*)
Pseudotrebouxia jamesii
(*Schaereria tenebrosa*)
Pseudotrebouxia potteri
(*Pertusaria* sp.
UTEX 900; *Rhizoplaca chrysoleuca*)
Pseudotrebouxia showmanii
(*Lecanora hageni*)
Pseudotrebouxia usneae
(*Usnea filipendula*)
Pseudotrebouxia sp.
(*Aspicilia calcarea*
Astroplaca sp.
Caloplaca holocarpa
Diploschistes scruposus
Lecania sp.
Lecanora tenera
Lecidea crystallifera
Lecidea sarcogynoides
Omphalodium arizonicum
Parmelia tinctorum
Rhizocarpon geographicum
Toninia caerulonigricans
Xanthoria parietina)
*Radiosphaera dissecta**
(UTEX 1187; non-lichenized)
*Spongiochloris excentrica**
(UTEX 108; non-lichenized)
*Spongiochloris llanoensis**
(UTEX 1245; non-lichenized)
Trebouxia anticipata
(UTEX 903; *Parmelia rudecta*)
Trebouxia crenulata
(CCAP 219/lb; *Parmelia*
acetabulum)
CCAP 219/2; *Xanthoria*
aureola)
Trebouxia gelatinosa
(UTEX 905; *Parmelia caperata*)

Algal and fungal symbionts were separated into culture using standard isolation techniques. The algal symbionts were grown on Bold's basal medium (BBM)x3N + 1% glucose agar slants. Single spore isolates of the fungal symbiont were cultured in Lilly and Barnett liquid medium or malt–yeast extract broth liquid medium[13]. The fungal cultures were centrifuged at 10,000 r.p.m., washed with distilled water and recentrifuged at the same speed. Clumps of hyphae were taken from the pellet and mixed with different algae. These symbiont mixtures were inoculated onto the surface of newly cleaved mica strips (1 × 7.5 cm) which had been soaked in BBMx3N solution. Each strip was located in a 125-ml Erlenmeyer flask; one end of the strip was embedded in 20 ml of plain 2% agar and the other end rested against the flask. The flask was plugged with non-absorbent cotton covered with aluminium foil. The relative humidity in the flask was about 95%. Incubation of the flasks was at 3000 lx, 12/12 light–dark cycle, at 21 °C during the light period and 19 °C in the dark period. Scanning electron microscope observations were made 2 months after inoculation.

* Mycobiont used in these crosses was isolate no. 6.

Haustoria were common in the natural and synthetic lichens (Fig. 2*a*), but in the former, about twice the proportion of algal cells was penetrated than in the latter (Table 2). Algae in the newly synthesized squamules were probably dividing more rapidly than those in the natural thalli, thus being less susceptible to fungal contacts.

Algae that were incompatible with the fungus (Table 1) were parasitized and destroyed (Fig. 2*b*). These included many green and one blue-green phycobionts and non-lichenized forms. Because initial binding of the fungal hyphae to algal cells seems to be similar in all cases, compatibility must be a secondary phenomenon. Perhaps there is algal resistance to fungal parasitism; that is, some algae may succumb more slowly than others to particular fungi.

In *C. cristatella*, the relationship between fungus and alga seems to be controlled parasitism. In the natural lichen, although more than 50% of the phycobiont cells were penetrated by haustoria, the algal population seemed healthy. In our successful laboratory syntheses, we saw algal cells apparently killed by haustoria and hyphae growing through them. Therefore, any resistance that compatible algae may have to the fungus is not permanent. In nature, a balanced relationship may have evolved between the symbionts. It is possible that the parasitism of the natural phycobiont is balanced by proliferation of the algal cells. Others have suggested controlled parasitism[5-7] and it may be a general feature of lichens.

The role of the haustoria is not clear. Evidence indicates that they do not transfer metabolites from alga to fungus[8,9]. They might be vestigial structures of ancestral fungi that were more actively pathogenic on other plants. A relationship between lichen fungi and unlichenized fungal parasites is suggested by the concentric bodies, cellular organelles of unknown function found in numerous lichen fungi and ascomycete plant pathogens.

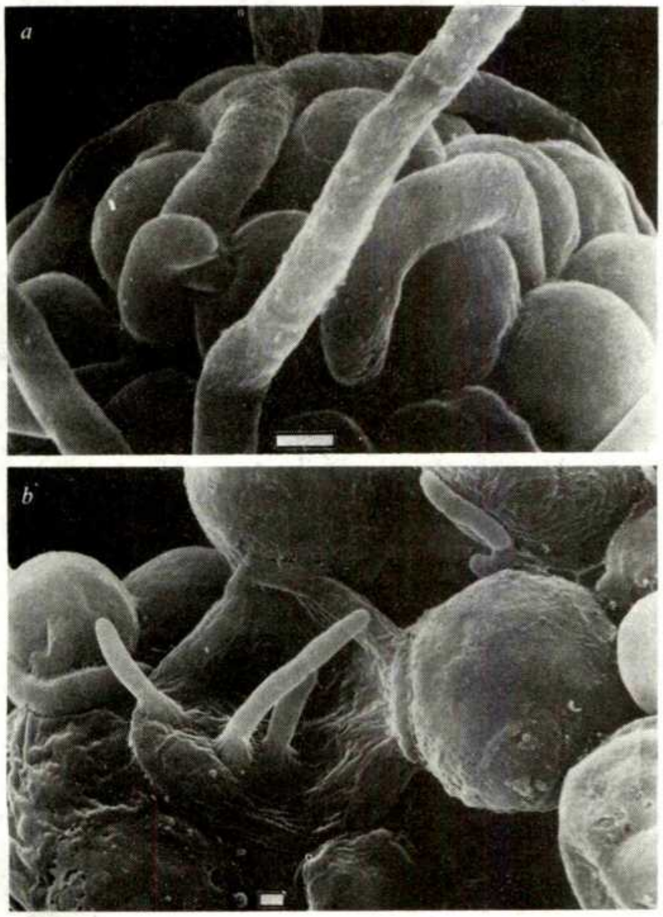


Fig. 2 Fungal intrusions into algal cells. Procedures used were as described in Fig. 1 legend. *a*, Cell of *Trebouxia erici* showing invaginated wall around haustorium. *b*, Three hyphae growing out of collapsed cell of *Trebouxia gelatinosa*. Scale bars, 1 μm.

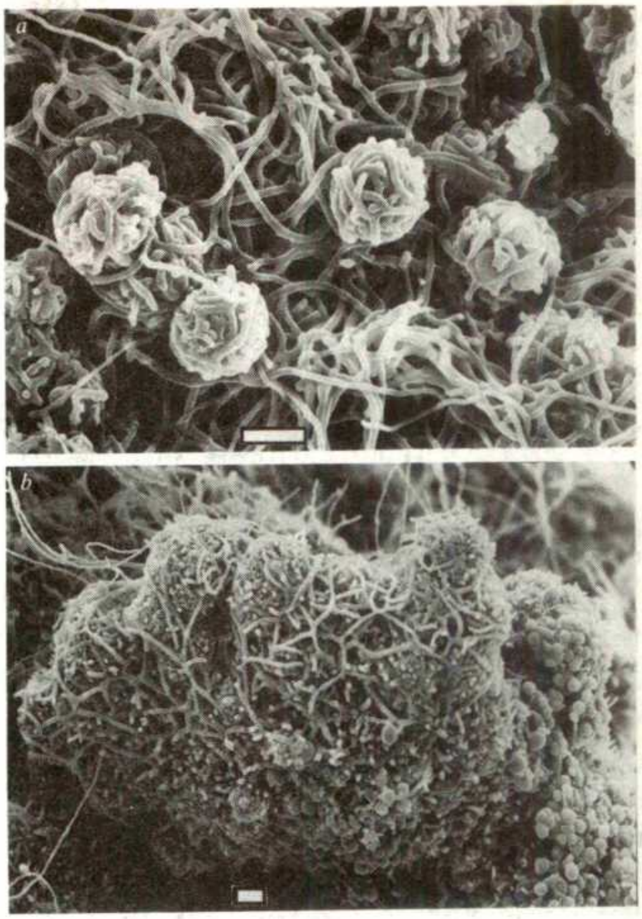


Fig. 3 Development of lichen squamules. See Fig. 1 legend for procedures. *a*, Presquamules formed by algal cells enveloped by mycobiont. *b*, Presquamules linked together by a network of superficial hyphae. Scale bars, 10 μm.

Some algal cells in the early stages of the successful syntheses were almost completely covered by fungal hyphae, which may be a response to the nutritional needs of the fungus. The mica substrate contained only mineral supplements; thus, the algal cells provided the only source of organic nutrients for the fungus. Farrar[10] estimated that 90% of the carbon fixed in photosynthesis by natural lichens is released by the algae but this release virtually ceases when the algae are removed from the

Table 2 Frequency of haustoria in natural and resynthesized squamules of *Cladonia cristatella*

Natural lichen	Total cells observed	Cells penetrated by haustoria	% Cells penetrated
Squamule 1	200	113	57
Squamule 2	100	65	65
Squamule 3	184	107	58
Resynthesized lichens			
C. cristatella s.s.8 × *Trebouxia excentrica* (*Cladonia leporina*)	200	54	27
C. cristatella s.s.8 × *Trebouxia erici*	100	24	24

Squamules were placed in a drop of Loeffler's methylene blue or lactophenol cotton blue and gently crushed with a cover slip using sufficient pressure to rupture vegetative cells. The ruptured cells extruded their contents and left only the empty shell. Haustoria adhered strongly to the algal wall and were not dislodged by the crushing. Haustoria were readily visible because there were no algal contents to mask their presence and they stained whereas the algal wall remained unstained. Empty algal cells over 5 μm were counted randomly during transects from one side of the slide to the other. Crushed cells were not counted.

symbiosis. The algae we used were growing separately in culture and presumably released little or no carbohydrate. The initial extensive contacts of algal cells by fungal hyphae may be necessary to re-establish permeability of the algal cells. We did not find an extensive hyphal covering of algal cells in the natural lichen. Collins and Farrar[9] reported that in *Xanthoria parietina* only 20% of the algal wall was in contact with fungal walls. This minimal fungal coverage may be related to the large release of organic material by the alga in the symbiotic condition.

The lichenization of the free-living alga *F. israeliensis* proceeded only to the presquamule stage. The mixed culture showed numerous patches of dead algal cells killed by fungal intrusions. Cells of *P. terrestre* were enveloped by hyphae but presquamules did not form because the algae could not survive association with the fungus. *Trebouxia* belongs to the order Chlorococcales. *Friedmannia*, like *Pseudotrebouxia*, is in the Chlorosarcinales because it has vegetative cell division and *Pleurastrum* is in the Chaetophorales. Although they occupy divergent taxonomic positions, these algae have similar zoospores and division stages of zoosporogenesis[11,12].

Our scanning electron microscope observations of the natural lichen showed many bacteria on the surface of the squamules. However, the axenic development of squamules in our cultures revealed that bacteria were not involved in the morphogenetic development of this lichen.

Syntheses took place in constant humidity (about 95%) in a closed flask. Other than the initial drying after the wet fungus was placed onto the mica strip, the culture did not dry further; thus, the lichen developed despite the absence of drying and wetting cycles thought to be essential in lichen synthesis.

Most chemical substances of natural lichens are products of the symbiosis. Analyses (TLC and ion-HPLC) of our synthetic lichens revealed the presence of natural secondary products. *C. cristatella* s.s.6 × *Trebouxia erici* showed barbatic acid, 4-*O*-demethylbarbatic acid, obtusatic acid, norobtusatic acid, 3-α-hydroxybarbatic acid, didymic acid and rhodocladonic acid. *C. cristatella* s.s.6 × *F. israeliensis* showed barbatic acid, obtusatic acid and 4-*O*-demethylbarbatic acid.

This research was supported by a grant to V. A. from the NSF (PCM-7921676). We thank Paul Stein for technical assistance and discussions, and Chicita F. Culberson for TLC and HPLC analyses.

Received 15 August; accepted 11 November 1980.

1. Smith, D. C. in *Cellular Interactions in Symbiosis and Parasitism* (eds Cook, C. B., Pappas, P. W. & Rudolph, E. D.) 197–227 (Ohio State University Press, Columbus, 1980).
2. Schwendener, S. *Bot. Ztg.* **26,** 289–292 (1868).
3. Fink, B. *Mycologia* **5,** 97–166 (1913).
4. Ahmadjian, V., Jacobs, J. B. & Russell, L. A. *Science* **200,** 1062–1064 (1978).
5. Peirce, G. T. *Am. Nat.* **34,** 245–253 (1900).
6. Kershaw, K. A. & Millbank, J. W. *New Phytol.* **69,** 75–79 (1970).
7. Webber, M. M. & Webber, P. J. *Can. J. Bot.* **48,** 1521–1524 (1970).
8. Hessler, R. & Peveling, E. *Z. Pflanzenphysiol.* **86,** 287–302 (1978).
9. Collins, C. R. & Farrar, J. F. *New phytol.* **81,** 71–83 (1978).
10. Farrar, J. F. in *CRC Handbook Series in Nutrition and Food* B (in the press).
11. Molnar, K. E., Stewart, K. D. & Mattox, K. R. *J. Phycol.* **11,** 287–296 (1975).
12. Deason, T. R., Ryals, P. E., O'Kelley, J. C. & Bullock, K. W. *J. Phycol.* **15,** 452–457 (1979).
13. Ahmadjian, V. *The Lichen Symbiosis* (Ginn & Co., Lexington, Massachusetts, 1967).

Light–shade adaptation of *Stylophora pistillata*, a hermatypic coral from the Gulf of Eilat

Paul G. Falkowski

Oceanographic Sciences Division, Brookhaven National Laboratory, Upton, New York 11973

Zvy Dubinsky

Department of Life Sciences, Bar Ilan University, Ramat Gan, Israel

All reef-forming, or hermatypic, corals harbour photosynthetic endosymbiotic algae called zooxanthellae[1–5], which are assumed to be predominantly a single dinoflagellate species, *Gymnodinium microadriaticum* Freudenthal[6]. The zooxanthellae are essential for the well-being of their hosts[7–9]; nevertheless, little is known about how light affects the symbiotic association, especially regarding the numbers of zooxanthellae, their photosynthetic responses, and their overall productivity[10–14]. On the reefs of the Gulf of Eilat, *Stylophora pistillata* is an abundant hermatypic coral[15]; it is unique in that region in that it can adapt to a wide range of light intensities. In the high light intensities of lagoons or the upper areas of reefs, the corals are markedly lighter in colour than those living under ledges, in grottos, or near the reef floor (~15 m; Fig. 1). We report here on the biochemical and physiological adaptations of *S. pistillata* to variations in light intensity spanning more than two orders of magnitude.

Specimens of *S. pistillata*, adapted to light and shade regimes, were collected from a remote reef near Nabq in the Sinai Peninsula (28.20° N, 034.95° E) and taken immediately to a lagoon near a shore-based laboratory. The corals were kept in shaded (~10% incident light), submerged baskets in the lagoon until used. Within 2 h of collection, samples were weighed and photographed against a calibrated grid, and chlorophylls were extracted by freezing in liquid N_2 and thawing in 90% acetone on a reciprocating shaker under dim light for 1 h. Extraction was considered complete when little or no chlorophyll fluorescence was observed on the coral branches under long wavelength UV light[16]. The absorption spectra of the acetone extracts were scanned from 400 to 750 nm on a Beckman DB-G grating spectrophotometer and chlorophylls *a* and *c* were calculated according to equations of Jeffrey and Humphrey[17]. Soft tissue protein, extracted in boiling NH_3OH, was measured by the method of Lowry[18] using bovine serum albumin as a standard. To determine coral surface areas, photographs were made of individual branches (of known weight) against a calibrated grid and projected onto graph paper. The two-dimensional surface areas were obtained by tracing the coral branches and weighing the paper cutouts of the branches. These two-dimensional areas (that is, cross-sectional areas) were multiplied by π to estimate three-dimensional surface areas, assuming that the coral branches approximate a series of cylinders.

Zooxanthellae were removed from branches with a Water Pik (Teledyne)[19], suspended in fresh seawater, and counted by light

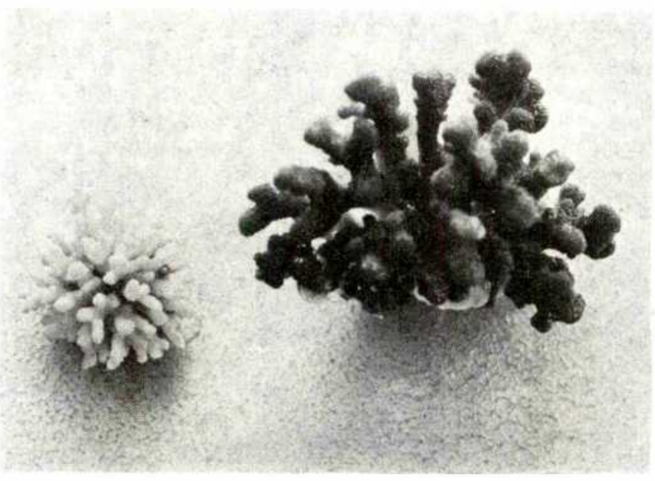


Fig. 1 Photograph of light-adapted (left) and shade-adapted specimens of *Stylophora pistillata* collected from the coral reefs in the Gulf of Eilat.

Table 1 Characteristics of pigment and protein composition of light-adapted, shade-adapted and transplanted individuals of *S. pistillata* and zooxanthellae

	Whole-coral characteristics						Zooxanthellae characteristics		
Coral type	*% I_0	*n*	†g cm^{-2}	Protein (mg cm^{-2})	Chl *a* (μg cm^{-2})	Cell density (cells cm^{-2}) ×10^6	Chl *a* (pg per cell)	Chl *a*/P_{700} (Molar ratio)	P_{700} content (pmol per cell × 10^{-6})
Shade-adapted	0.3–1.0	10	2.2±0.3	1.9±0.3	14.2±4.0	1.7±0.3	8.3±0.5	1,650±175	5.7±0.8
Light-adapted	50–90	11	4.1±0.6	1.8±0.6	3.6±1.1	1.6±0.3	2.2±0.3	425±50	5.9±0.5
Ratio (shade/light)	~200		1.9	1.0	3.9	1.1	3.7	3.9	1.0
Light-to-shade transplant	From 50–90 To 0.3–1.0	3	4.1±0.6	1.9±0.6	11.2±5.1	1.6±0.2	6.9±1.2	1,250±350	6.2±1.0
Ratio (L → S transplant/light)	~200		1.0	1.0	3.1	1.0	3.1	2.9	1.1
Shade-to-light transplant	From 0.3–1.0 To 50–90	3	2.2±0.3	1.9±0.5	3.2±0.9	1.4±0.3	2.3±0.5	450±75	5.7±0.6
Ratio (S → L transplant/light)	~1		1.9	1.0	0.9	0.9	1.0	1.1	1.0

The values for transplanted corals are given after 30 days of adaptation to the new light regime. (Values reported are means ±s.e. where applicable.)

* Range of the per cent of the surface light intensity.

† Surface area to mass ratio.

microscopy with a haemocytometer. Suspensions were filtered onto Whatman GF/C glass fibre filters; chlorophylls were extracted by grinding the filters in 90% acetone and quantified spectrophotometrically.

To determine photosynthetic responses, coral branches were placed in 250-ml Wheaton bottles in a constant temperature bath with a self-stirring oxygen electrode (Yellow Springs Instruments, model 5420A) connected to a 356H operational amplifier. Light was provided by a 500 W quartz halogen source and attenuated with Wratten neutral density filters. To provide more even illumination, a mirror was placed directly behind the branches. The size and number of photosynthetic units (PSUs) were estimated by measuring chlorophyll *a*/P_{700} ratios with an Aminco DW-2a spectrophotometer[20,21]. Light intensities were measured *in situ* with a Lambda LI 185 quantum meter interfaced with a 192S underwater sensor.

S. pistillata can be divided into two major morphological groups: light- and shade-adapted individuals (Table 1). The former have thinner and more numerous branches (Fig. 1), giving a ratio of surface area/weight 1.9 times greater than that of shade-adapted corals. In most cases shade-adapted colonies are subspherical, and only rarely have we observed the flattening previously described[22,23]. Protein/surface area ratios were remarkably constant for all individuals, however, and it is inferred that ratios of protein per gramme are directly proportional to surface areas per gramme for a given coral species.

Shade-adapted corals contained 7.4 times more chlorophyll *a* per gramme (3.9 times more per square centimetre) than those adapted to high light intensities. Differences in pigmentation reflect the intracellular pigment content of the zooxanthellae and cannot be attributed to differences in numbers of zooxanthellae per surface area of coral (Table 1). Those differences were primarily associated with changes in the number of chlorophyll *a* molecules per photosystem I reaction centre, P_{700}. The number of reaction centres was virtually constant in light- and shade-adapted zooxanthellae, suggesting that adaptation is mediated by changes in the size, rather than the number of PSUs, as has been observed in free-living unicellular algae[24,25]. The range of sizes of PSUs observed in *G. microadriaticum* (approximately 400–1,800 chlorophyll per P_{700}) is similar to those found in free-living dinoflagellates[24,25], but is larger than for chlorophytes (approximately 450 chlorophyll per P_{700})[25] and higher plants (approximately 350–500 chlorophyll per P_{700})[26]. Such changes were associated with changes in the photosynthetic responses to changes in irradiance (that is, P–I curves, Fig. 2). On the basis of coral surface area, light-adapted specimens had lower photosynthetic capacities and higher compensation light intensities than shade-adapted individuals (Fig. 2*a*). Instantaneous light-saturated net photosynthetic O_2 production by zooxanthellae exceeded coral respiratory demands in both light- and shade-adapted specimens. In neither was photoinhibition observed, even at intensities approaching 2,100 μeinstein m^{-2} s^{-1} (full noon sunlight in the summer in the surface waters of the Gulf of Eilat). When P–I curves are calculated on the basis of chlorophyll content (Fig. 2*b*), maximum photosynthetic capacities were not significantly different in light- or shade-adapted corals. However, shade-adapted corals appeared to utilize light more efficiently, as indicated by steeper initial

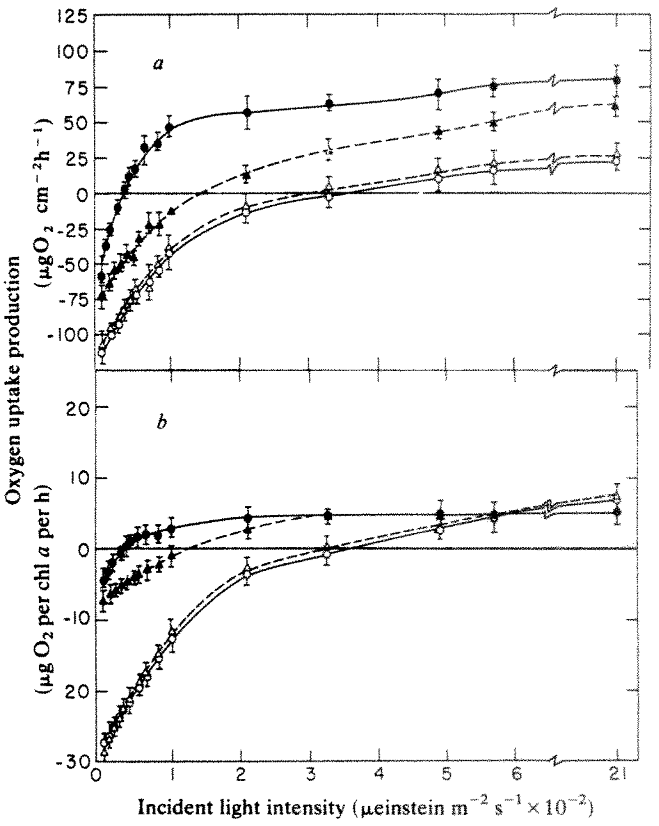


Fig. 2 Photosynthesis–irradiance profiles for light-adapted, shade-adapted, and transplanted *S. pistillata*. *a*, Light curves expressed on a surface area basis: ○, light-adapted; ●, shade-adapted; ▲, light-to-dark transplant after adaptation for 30 days; △, dark-to-light transplant after adaptation for 30 days. *b*, Light curves expressed on a chlorophyll *a* basis. The symbols used are the same as in *a*. Data points represent means and standard deviations ($n = 3$ for light adapted, $n = 3$ for shade adapted, $n = 2$ for dark-to-light transplants, $n = 3$ for light-to-dark transplants). All measurements were made at 28 ± 1 °C.

slopes of the P–I curves. These P–I responses were qualitatively similar to those observed in the foliaceous coral *Pavona praetorta* Dana from Enewetak Atoll[12] and are consistent with theoretical arguments that changes in the size of PSUs ought to affect the efficiency of light utilization[21,24,25].

Corals were able to adapt when transplanted from one light regime to another (Table 1, Fig. 2). The process was quicker for transplantation from lower to higher light intensities, being effectively complete after 4 weeks. Corals transplanted from high to low light intensities were not fully shade adapted after 4 weeks, but continued to adapt slowly for 6 to 8 weeks. During that period, light-to-shade transplant pigmentation increased in those corals and P–I responses were intermediate between the two extremes. The difference in adaptation time between the two types of transplants may be related to temperature in the two environments[16,27].

Previous attempts to transplant corals in this way have not generally been successful as the corals often show signs of physiological stress. This has been interpreted as indicating that light- and shade-adapted corals represent locally adapted populations or ecotypes[23]. Usually a particular species appears to be associated with an optimum light regime, and changes in light intensity result in changes in species composition[28]. Our reciprocal transplant results strongly suggest, however, that changes in coral zonation associated with a gradient in light intensity are not necessarily caused by the inability of zooxanthellae to adapt to wide variations in light regimes. It would seem that in the case of *S. pistillata* the wide range of light intensities tolerated by this coral is not due to zooxanthellae populations of distinct ecotypes as described for the West Indian coral *Montastrea annularis*[23]. In particular, the ability of *S. pistillata* to adapt to various light regimes suggests that light- and shade-adapted corals are not always ecotypes and that the niche width of a coral species with respect to light *per se* is not limited by its zooxanthellae.

This research was supported by the US-Israel Binational Science Foundation and assisted, in part, by the US Department of Energy under contract DE-ACO2-76CH00016. We thank Jacob Schafer, Ruth Benaksas, David Sherf and Tova Sa'adon for technical assistance, and Dr M. Avron for the use of the Aminco spectrophotometer.

Received 7 May; accepted 22 September 1980.

1. Droop, M. R. *Symp. Soc. gen. Microbiol.* **13,** 171–199 (Cambridge University Press, London, 1963).
2. Taylor, D. L. *A. Rev. Microbiol.* **27,** 171 (1973).
3. Goreau, T. F. in *The Biology of Hydra and of Some Other Coelenterates* (eds Lenhoff, H. M. & Loomis, W. F.) 269–285 (University of Miami Press, 1961).
4. Taylor, D. L. in *Biology and Geology of Coral Reefs* (eds Jones, O. A. & Endean, R.) 267–277 (Academic, New York, 1973).
5. Odum, H. T. & Odum, E. P. *Ecol. Monogr.* **25,** 291 (1955).
6. Yonge, C. M. *Bull. mar. Sci.* **23,** 1 (1973).
7. Pearse, V. B. & Muscatine, L. *Biol. Bull.* **141,** 350 (1971).
8. Barnes, D. J. & Taylor, D. L. *Helgolander wiss. Meersunters* **24,** 284 (1973).
9. Muscatine, L., Pool, R. R. & Cernichiari, E. *Mar. Biol.* **13,** 298 (1972).
10. Wethey, D. S. & Porter, J. W. *Nature* **262,** 281 (1976).
11. Kawaguti, S. & Nakayama, T. *Biol. J. Okayama Univ.* **16,** 67 (1973).
12. Drew, E. A. *J. exp. mar. Biol. Ecol.* **13,** 165 (1973).
13. Zvalinski, V. I., Titlyanov, E. A., Lelefkin, U. A. & Novozhilov, A. V. in *Biology of Coral Reefs: Photosynthesis of Zooxanthellae and Macroalgae Acad. Sci. U.S.S.R. Trans.* (Russian) Vol. 12 (ed. Zvalinsky, V. I.) 29–52 (Vladivostok, 1978).
14. Muscatine, L. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 381–402 (Plenum, New York, 1980).
15. Loya, Y. *Nature* **259,** 478 (1976).
16. Riper, D. M., Owens, T. G. & Falkowski, P. G. *Pl. Physiol.* **64,** 49 (1979).
17. Jeffrey, S. W. & Humphrey, G. F. *Biochem. Physiol. Pflanz.* **167,** 191 (1975).
18. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193,** 265 (1951).
19. Johannes, R. E. & Weibe, W. J. *Limnol. Oceanogr.* **15,** 822 (1970).
20. Marsho, T. U. & Kok, B. *Meth. Enzym.* **23,** 515 (1971).
21. Falkowski, P. G. & Owens, T. G. *Pl. Physiol.* **66,** 592 (1980).
22. Barnes, D. J. *Bull mar. Sci.* **23,** 280 (1973).
23. Dustan, P. *Bull mar. Sci.* **29,** 79 (1979).
24. Prezelin, B. B. & Alberte, R. S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1801 (1978).
25. Falkowski, P. G. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 99–120 (Plenum, New York, 1980).
26. Boardman, N. K. *A. Rev. Pl. Physiol.* **28,** 355 (1977).
27. Hitchcock, G. L. thesis, Univ. Rhode Island (1977).
28. Chalker, B. E. & Taylor, D. L. *Proc. R. Soc.* **B190,** 323 (1975).

Heterozygous advantage expressed through sexual selection in a polymorphic African butterfly

D. A. S. Smith

Department of Biology, Eton College, Windsor,
Berkshire SL4 6EW, UK*, and Department of Life Sciences,
University College at Buckingham, Buckingham, UK

A field study of the polymorphic African butterfly *Danaus chrysippus* L. at Dar es Salaam, Tanzania (1972–73) indicated that mating may be non-random between forms *dorippus* and *aegyptius*[1] and that seasonal sexual selection occurs[2]. Subsequently, a long-term breeding programme revealed a heterozygous phenotype[3] which closely resembles *dorippus*, and allowed estimates of heterozygote frequency in the field[4]. Analysis of mating records (1974–75) shows that males visibly heterozygous at the controlling C locus (from *transiens*)[3,4] possessed a highly significant mating advantage over both homozygous phenotypes throughout the study period, although pair composition among the three genotypes was random. I propose here that balancing selection, mediated mainly through male competition for mates, makes a decisive contribution to the control of the polymorphism.

Sampling was carried out several times a week for 17 consecutive months (April 1974–August 1975). All captures were marked before release to avoid double counting. The data (Table 1) consist of 340 observed pairings and 1,519 unmated males. As monthly samples were too small to be of much value, they were amalgamated into two annual seasons, one being verdant and relatively cloudy, wet and cool (April–August), when the frequency of *aegyptius* rises rapidly, and the other hot and dry punctuated by a short rainy season (September–March), when *dorippus* predominates[5]. Heterogeneity between mated and unmated samples was tested in $2 \times 3 \times 3$ contingency tables and χ^2 analysed[6] (Table 2).

The analysis shows that *transiens* males had consistently and very significantly higher success in mating than both homozygous phenotypes and that the effect was independent of season and changes in morph-ratio. Residual heterogeneity was not significant. As *CC* and *Cc* butterflies are not always recognizable phenotypically, genotype numbers (Table 3) are estimated by applying a correction for expressivity based on breeding data. The expressivity of *c* in known heterozygotes is modified by a closely linked B locus: 75% $(n = 173)$ of *B-Cc* and 45% $(n = 129)$ of *bbCc* butterflies were *transiens*[4], and these values were used to estimate *Cc* numbers from the number of *transiens* observed. Fitness (Table 3) of *CC* males was low (9–18%) in both wet seasons, rising to 79% in the dry season. In contrast, *cc* males had a fitness of 60–64% in the wet seasons, falling to 15% in the dry season. Differential and seasonally alternating mating fitnesses of this magnitude must make a major contribution to the synchronized cycling in morph frequency observed in four successive years (1972–75)[4,5]. Within seasons, there is no evidence for non-random mating among the three C-locus genotypes (within the mating samples) or for sexual selection in females.

How could such powerful selection for heterozygous males operate? Female preference for them could contribute but is probably not the key factor because it would imply synchronized seasonal switches of the secondary preference (for either *CC* or *cc* males) by all three female genotypes—if this were not the case, non-random mating would be detectable in large samples. *Cc* males (but not females) are significantly larger than both

---

* Address for correspondence.

0028–0836/81/030174—02$01.00

Table 1 Numbers of C-locus phenotypes of male *Danaus chrysippus* caught mating and unmated at Dar es Salaam in 1974–75

Period	Condition	Phenotypes			Total	χ^2_2
		dorippus	*transiens*	*aegyptius*		
April–August 1974	Mated	36	52	23	111	11.297
	Unmated	269	187	155	611	$P = 0.0035$
	Total	305	239	178	722	
September 1974–March 1975	Mated	88	47	2	137	13.451
	Unmated	292	113	47	452	$P = 0.0012$
	Total	380	160	49	589	
April–August 1975	Mated	26	40	26	92	12.683
	Unmated	194	117	145	456	$P = 0.0018$
	Total	220	157	171	548	
Overall	Mated	150	139	51	340	26.519
	Unmated	755	417	347	1,519	$P = 1.74 \times 10^{-6}$
	Total	905	556	398	1,859	

homozygotes[4], which could give them a competitive edge through faster flight. However, the mean forewing length of mated males ($\bar{x} = 37.9$ mm, $n = 161$) did not differ significantly from those unmated ($\bar{x} = 38.0$ mm, $n = 881$) between October 1974 and September 1975; thus selection based on size alone is ruled out. Alternatively, the male aphrodisiac pheromone, essential in courtship[7–9], might differ qualitatively or quantitatively between genotypes, although this possibility received no support from a chemical analysis of 50 pairs of hairpencils and wing pouches from each genotype (J. Meinwald, personal communication). However, it is known that male *D. chrysippus* must spend several days collecting pheromone precursors (pyrrolizidine alkaloids) from plants; therefore those males able to maintain the highest foraging rates are likely to achieve earlier matings and greater overall success.

The major colour differences between the genotypes (see ref. 4, Figs 1–16) could influence metabolic rate and hence activity levels. The amount of black and brown pigment is greatest in *aegyptius* and least in *dorippus* with most *transiens* intermediate. The darker coloration of *aegyptius* could allow it to maintain a higher body temperature, due to superior absorbance of radiant heat, during the cooler, cloudier months. It is interesting that *aegyptius* is the only form in the cooler subtropical parts of the species' range at both its northern and southern limits. Conversely, enhanced reflectance could provide relative protection for *dorippus* during the hot, dry season, a suggestion supported by the fact that it is the only form in the arid regions of northern Kenya and Somalia[4,10]. The intermediate heterozygote could be better adapted than either homozygote to median radiation values and therefore better able to maintain optimum activity through the vicissitudes of a climate which can feature substantial wet or dry periods out of season, with successive years rarely having a consistent pattern.

Table 2 Analysis of χ^2

Source of variation	χ^2	Degrees of freedom	Probability
Morphs × periods	124.16842	4	9.00×10^{-10}
Proportions mating × periods	14.66944	2	6.52×10^{-4}
Proportions mating × morphs	26.51889	2	1.74×10^{-6}
Residual heterogeneity	6.72306	4	0.151
Total	172.07981	12	

Table 3 Fitness of C-locus homozygotes relative to the heterozygote

Period	Condition	Genotypes			Total
		CC	*Cc*	*cc*	
April–August 1974	Mated	3	85	23	111
	Unmated	149	307	155	611
	Fitness	*0.091*	*1*	*0.596*	
September 1974–March 1975	Mated	59	76	2	137
	Unmated	209	196	47	452
	Fitness	*0.788*	*1*	*0.146*	
April–August 1975	Mated	5	61	26	92
	Unmated	114	197	145	456
	Fitness	*0.178*	*1*	*0.643*	

Genotype numbers are based on corrections for expressivity (see text).

These results are of interest in three respects: (1) they provide a rare[11] and unusually explicit example of heterozygous advantage at a single locus in a wild population. (Linkage disequilibrium involving the C and two other loci[3-5] and epistasis[4] implies that the real unit of selection is a supergene or inversion.) (2) The data show that heterozygous advantage is a mechanism capable of both conserving and releasing as required the genetic diversity on which selection can act cyclically in an oscillating climate. I emphasize here that the generation time of *D. chrysippus* is short in equatorial regions with 13–14 overlapping generations a year and that similar marked changes in morph frequency are known in other tropical species such as *Catopsilia florella*[12] and *Hypolimnas misippus*[13,14], (3) I conclude that male competition predominantly determines the outcome of sexual selection in *D. chrysippus* and outweighs female choice, receptivity or allure which are more often identified as critical factors in insects[15].

I thank Professor J. Meinwald for pheromone analyses on male *D. chrysippus* from Dar es Salaam and Dr P. O'Donald for helpful suggestions.

Received 20 August; accepted 7 November 1980.

1. Smith, D. A. S. *Nature* 242, 129–130 (1973)
2. Smith, D. A. S. *Science* 187, 664–665 (1975)
3. Smith, D. A. S. *Ent. Scand.* 6, 134–144 (1975).
4. Smith, D. A. S. *Zool. J. Linn. Soc.* 69, 87–109 (1980)
5. Smith, D. A. S. *Heredity* 34, 363–371 (1975)
6. Lewis, B. N. *J. R. statist. Soc.* 125, 88–117 (1962).
7. Boppré, M. *Ent. exp. appl.* 24, 64–72 (1978)
8. Meinwald, J. *et al. Experientia* 30, 721–722 (1974)
9. Schneider, D. *et al. J. comp. Physiol.* 97, 245–256 (1975)
10. Pierre, J. *C. r. Somm. Séanc. Soc. Biogéogr.* 486, 179–187 (1980)
11. Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York, 1974).
12. Owen, D. F. *Tropical Butterflies* (Clarendon, Oxford, 1971).
13. Smith, D. A. S. *Biol. J. Linn. Soc.* 8, 183–204 (1976)
14. Edmunds, M. *Heredity* 24, 281–302 (1969)
15. O'Donald, P. *Genetic Models of Sexual Selection* (Cambridge University Press, 1980)

Fluoromevalonate acts as an inhibitor of insect juvenile hormone biosynthesis

Gary B. Quistad, David C. Cerf, David A. Schooley & Gerardus B. Staal

Departments of Biochemistry and Biology, Zoecon Corporation,
975 California Ave, Palo Alto, California 94304

The potential of analogues of insect juvenile hormones (JHs) for pest control has been demonstrated by the application of methoprene for the selective control of certain insects, including several mosquito species, the cigarette beetle and the hornfly[1]. Compounds that impede JH production could offer a potentially more attractive method of control because accelerated metamorphosis would shorten the lifetime of the larvae, the stage that usually does maximum damage to the host plant. Of the few anti-juvenile hormone (AJH) agents described[2–5], only precocenes evoke all the required biological responses, but they affect a few insect taxa of minor economic importance. Their activity is thought to result from cytotoxicity[6,7] in the corpora allata. In view of similarities between the biosynthesis of the carbon skeletons of insect JHs[8,9] and farnesyl pyrophosphate (a precursor of steroids and other terpenoids[10]), compounds that inhibit the early steps in the biosynthesis of cholesterol might be expected to suppress JHs. Although this concept has already been explored without success[11,12], we now report the inhibition of juvenile hormone biosynthesis (AJH activity, see Table 1 legend) by a fluorinated mevalonate (tetrahydro-4-fluoromethyl-4-hydroxy-2H-pyran-2-one, FMev, Fig. 1) which until now was known only to be hypocholesterolaemic[13–15].

AJH activity was observed initially when 0–24-h third-instar larvae of the tobacco hornworm (*Manduca sexta*) were treated topically with 100 μg of FMev, which induced three types of effects, all directly related to JH deficiency. First, intense black pigmentation was often observed after the succeeding moult (to fourth instars). In *Manduca* this is known to result from JH deficiency[16,17] and it therefore serves as a diagnostic tool in screening for JH suppressing activity. There were further signs of antagonism at the end of this instar, when many larvae started to show the prepupal burrowing behaviour normally occurring one instar later. In this case, the subsequent moult (to the fifth instar) often resulted in miniature pupae or intermediates between larvae and pupae with a varying mosaic of larval and pupal cuticle (Fig. 2). These morphologically premature pupae and intermediates were generally nonviable. Given higher doses, third-instar larvae started prepupal burrowing at the end of the third instar and produced pupal intermediates in the fourth instar. All these effects of FMev could be averted by concurrent application of a very active juvenoid such as hydroprene; this 'rescue' confirms that the observed symptoms can be attributed to JH deficiency. A high dose of FMev (2.5 mg) injected into adult *M. sexta* 1–2 h after eclosion also prevented the maturation of eggs. This anti-gonadotropic activity is consistent with the expected properties of an AJH compound.

A considerable morphogenetic effect was observed after treatment of fifth-instar *Galleria mellonella* larvae with FMev. Untreated larvae normally developed through three additional instars before forming pupae approximately 15 mm long. Larvae treated with FMev produced premature pupae in any of these three subsequent instars, the smallest sixth-instar pupae being 4 mm long. Some of the large premature pupae survived to generate precocious miniature adults of otherwise normal appearance.

The AJH properties of FMev seem to be manifest only in lepidoptera. Larvae of the six moth species examined responded in various degrees to the AJH effects of FMev (Table 1). No such effects were found when the larval stages of the following non-lepidopteran insects were treated with FMev: *Musca domestica* (housefly), *Aedes aegypti* (yellow fever mosquito), *Schistocerca nitens* (vagrant grasshopper), *Dermestes maculatus* (hide beetle), *Hypera postica* (alfalfa weevil) and *Oncopeltus fasciatus* (milkweed bug).

Table 1 Anti-juvenile hormone* activity of FMev on several lepidopteran species

Insect	Instar treated	ED_{50} (mg per g)
Manduca sexta (tobacco hornworm)	Third	0.7
Samia cynthia (Cynthia moth)	Third	1.0
Phryganidia californica (California oak moth)	Third	1.0
Galleria mellonella (greater wax moth)	Fifth	40
Spodoptera exigua (beet armyworm)	Third	44
Heliothis virescens (tobacco budworm)	Third	>200

0–24-H larvae of the indicated instar of each species were treated topically with 1 μl acetone dilutions of FMev and reared to the last larval instar. Observations were made throughout the bioassays. The ED_{50} (effective dose causing 50% of the treated larvae not showing acute toxicity from FMev to demonstrate precocious metamorphosis or prepupal behaviour) is based on the mg mass of compound per g of insect. In the case of *H. virescens*, the ED_{50} was never achieved, although a small percentage of the treated larvae did manifest premature pupation.

* We use the term 'anti-juvenile hormone (AJH) agents' to describe those chemicals which induce symptoms of JH deficiency, including premature metamorphosis. This general term can encompass the precocenes (cytotoxic to the corpora allata[6,7]), FMev (inhibitor of JH biosynthesis) and AJH agents of as yet obscure[3–5] mechanism of action.

FMev exhibits moderate toxicity at high doses for lepidoptera which are less sensitive to the AJH effects of FMev. However, for more sensitive species such as *M. sexta*, toxicity is minor (5–10% at ED_{50}). Because this acute toxicity also occurs in insect species not responding with precocious metamorphosis and because concurrent application of a juvenoid did not rescue this mortality, we conclude that this toxicity is not related to inhibition of JH biosynthesis and results from another mechanism. A possible explanation for the nonspecific mortality is metabolic conversion of FMev to fluoroacetate, a potent biocide. Treatment of *M. sexta* larvae with ethyl fluoroacetate revealed substantial toxicity without AJH effects.

We originally tested FMev for AJH activity based on the rationale that it could serve as an antimetabolite of mevalonate and/or homomevalonate. Preliminary studies affirm that FMev

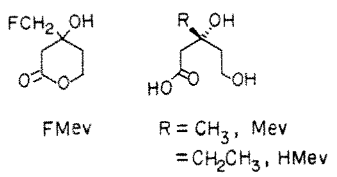


Fig. 1 Fluoromevalonolactone (FMev), mevalonate (Mev, precursor of cholesterol and insect juvenile hormones) and homomevalonate (HMev, biosynthetic precursor of several insect juvenile hormones). FMev has been synthesized previously[14,15], but we synthesized it by treating ethyl fluoroacetate (highly toxic) with allylmagnesium bromide, followed by ozonolysis of the resultant diallylcarbinol from the Grignard reaction. The ozonide was reduced with sodium borohydride to give 3-fluoromethyl-1,3,5-pentanetriol, which was selectively oxidized to FMev with chromium trioxide in aqueous sulphuric acid (Jones reagent).

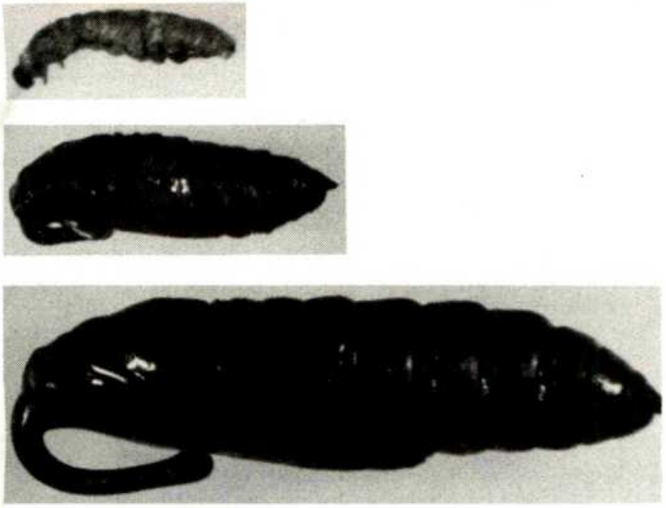


Fig. 2 A larval–pupal intermediate, precocious pupa and normal pupa (from top to bottom) for *Manduca sexta* (tobacco hornworm). The accelerated metamorphosis for the intermediate and precocious pupa was induced by treating third-instar larvae with 100 μg of FMev.

inhibits the biosynthesis of insect JHs. The titre of JHs in *M. sexta* haemolymph is markedly reduced when larvae are treated topically with the physiologically effective dose of FMev (J. P. Edwards and B. J. Bergot, unpublished). An *in vitro* study with homogenates of *M. sexta* corpora allata showed significant inhibition of metabolism of ^{14}C-mevalonate and ^{3}H-homomevalonate to subsequent precursors of JHs (F. C. Baker, unpublished).

We have synthesized many analogues of FMev, mevalonolactone and related compounds. Five of them show AJH activity on lepidopteran larvae, but are less active than FMev. Four of these marginally active compounds are likely to be active because of their facile metabolism to FMev (for example, 3-fluoromethyl-1,3,5-pentanetriol). The difluoromethyl analogue of FMev was five times less active on *M. sexta* then FMev, whereas the trifluoromethyl analogue was devoid of AJH properties.

We have thus found that a compound which is hypocholesterolaemic may also have AJH activity; perhaps other hypocholesterolaemic agents may serve as chemical structures which can be manipulated by analogue synthesis to optimize AJH activity. However, tests of several chemically unrelated materials with similar hypocholesterolaemic activity have revealed no other active inhibitors (ref. 12 and G.B.Q. *et al.* unpublished).

The activity of FMev by itself is not sufficiently high, even on lepidopteran pests, for commercial exploitation. Nevertheless, FMev should be an invaluable tool in studies of lepidopteran physiology and endocrinology. We hope the novel mode of action—inhibition of JH biosynthesis—can form the basis for new selective insect-control agents.

Received 23 June; accepted 16 October 1980.

1. Siddall, J. B. *Envir. Hlth Perspectives* **14,** 119–126 (1976).
2. Bowers, W. S. in *The Juvenile Hormones* (ed. Gilbert, L. I.) 394–408 (Plenum, New York, 1976).
3. Staal, G. B. *Scr. varia pont. Sci.* **41,** 353–383 (1977).
4. Murakoshi, S. *et al. Jap. J. appl. ent. Zool.* **19,** 267–272 (1975).
5. Slama, K. *Acta ent. bohem.* **75,** 65–82 (1978).
6. Brooks, G. T., Pratt, G. E. & Jennings, R. C. *Nature* **281,** 570–572 (1979).
7. Schooneveld, H. *Experientia* **35,** 363–364 (1979).
8. Schooley, D. A., Judy, K. J., Bergot, B. J., Hall, M. S. & Jennings, R. C. in *The Juvenile Hormones* (ed. Gilbert, L. I.) 101–117 (Plenum, New York, 1976).
9. Law, J. H. in *Polyisoprenoid Biosynthesis* (ed. Porter, J.) (in the press).
10. Richards, J. H. & Hendrickson, J. B. *The Biosynthesis of Steroids, Terpenes, and Acetogenins* (Benjamin, New York, 1964).
11. Matolcsy, G., Varjas, L. & Bordas, B. *Acta phytopath. hung.* **10,** 455–463 (1975).
12. Hammock, B. D. *et al. J. agric. Fd Chem.* **26,** 166–170 (1978).
13. Singer, F. M., Januszka, J. P. & Borman, A. *Proc. Soc. exp. Biol. Med.* **102,** 370–373 (1959).
14. Tschesche, R., Machleidt, H. & Bucher, T. US Patent 3,075,997, Jan. 29 (1963).
15. Tschesche, R. & Machleidt, H. *Justus Liebigs Annln Chem.* **631,** 61–76 (1960).
16. Safranek, L. & Riddiford, L. M. *J. Insect Physiol.* **21,** 1931–1938 (1975).
17. Staal, G. B. *Scr. varia pont. Acad. Sci.* **41,** 333–352 (1977).

A protein called immaturin controlling sexual immaturity in *Paramecium*

Nobuyuki Haga & Koichi Hiwatashi

Biological Institute, Tohoku University, Aoba-yama,
Sendai 980, Japan

As in many metazoans, clones of some species of *Paramecium* have, after conjugation, a period of immaturity during which the cells cannot mate. The duration of immaturity is measured by the number of cell generations, which remains fairly constant, although duration in time varies with rate of cell reproduction[1–3]. Genic involvement is shown by mutants with reduced periods of immaturity[4]. In three different groups of *Paramecium* species, the cytoplasm of immature cells apparently contains the same substance which represses mating activity when injected into sexually mature cells[5–7]. The immaturity-inducing substance seems to be absent from sexually mature cells, as brei made not only from mature cells in the stationary phase (mating-reactive cells), but also from those in the log phase (mating-non-reactive cells), does not repress mating activity when injected into mature cells[5]. Variations in the amount of the substance during immaturity suggest that it controls the duration of the period[8]. We have isolated and partially characterized the substance—a single protein called immaturin. The activity of immaturin is dose dependent and associated with a heat-labile protein of molecular weight (MW) 10,000.

For quantitative assay of immaturin activity by microinjection, we grew each injected cell in a glass capillary (0.6 mm inside diameter, 75 mm long) containing 20 μl fresh lettuce-juice medium inoculated with a non-pathogenic strain of *Klebsiella pneumoniae* 1 day before use[9]. The cell in each capillary produces exactly 16 cells which then enter synchronously into stationary phase. The 16 cells are isolated and then provided with mating-reactive cells of complementary mating type. If they are mature, at least 13 of the 16 form mating clumps with the cells of complementary type (each clump consisting of one cell of the type being tested, and a group of reactive cells of the opposite type). The activity of injected immaturin was measured by the percentage of sets of 16 forming less than 13 clumps (the clumps were usually from 10 to 0 when the injection was effective).

The immaturity period of the strains of *Paramecium caudatum* we used was about 50 fissions after conjugation. For isolation of immaturin, cells ($\sim 9 \times 10^7$), about 20 fissions after conjugation, were washed with solution A (10 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl, *p*H 7.4) and then with solution A′ (solution A + 0.25 M sucrose) by slow centrifugation (100–200*g*). The pellet was homogenized with an injection syringe[10] and the brei thus obtained was centrifuged at 105,000*g* for 60 min at 2 °C. The supernatant was filtered through a 0.45-μm Millipore filter. The filtrate was applied to a Sephadex G-50 column and each fraction was assayed by injecting 15-pl samples into mature cells. When the most active fraction (no. 32) was injected, 16 clones out of 20 formed less than 13 clumps when tested. Among those 16 clones, 4 formed no clumps and 10 formed less than 8 clumps. When the same fraction from mature cells was injected, 40 out of 51 formed 16 clumps and the remaining 11 formed more than 12 clumps. Then, the active fractions (nos 31, 32 and 33 in Fig. 1) were chromatographed on a DEAE-Sephadex A-25 column. The activity of immaturin was eluted as a single peak at about 400 mM NaCl (Figs 1, 2). In the test of the peak fraction (no. 10), 15 clones out of 19 injected formed less than 13 clumps. Among those 15, 10 formed less

0028-0836/81/030177—03$01.00

Table 1 Purification of immaturin

Step	Total protein obtained (mg)	No. of cells injected	Activity (%)	Protein injected per cell (pg)	Specific activity
(1) 105,000*g* supernatant	154.4	22	72.7	15	4.8
(2) Sephadex G-50 chromatography	3.6	25	88.0	0.7	125.7
(3) DEAE-Sephadex chromatography	0.013	25	80	0.03	2666.6

Activity is % of mature clones which contained more than 25% immature cells after injection. Specific activity is defined as activity per pg injected protein. Amount of protein was determined by the method of Lowry *et al.*[21].

than 9 clumps. The degree of purification was calculated on the basis of the activity of the 105,000*g* supernatant.

The relation between concentration of immaturin and activity of each concentration was first examined by injecting a dilution series of the supernatant. The activity (% of immaturity brought about by injection) wasnearly proportional to the amount of protein injected. When specific activity was defined as activity per pg injected protein, three purification steps increased the specific activity of immaturin about 550-fold (Table 1). To determine the purity of each fraction, samples from the DEAE fractions were subjected to acrylamide-gel electrophoresis by a modified method of Davis[11]. Lyophilized material from the peak fraction (no. 10) of the DEAE-Sephadex elution was dissolved in 300 μl of solution A and 100-μl samples were applied to a gel containing 11.25% acrylamide. Only a single band was strongly stained with Coomassie brilliant blue, and no minor band was detected (Fig. 3).

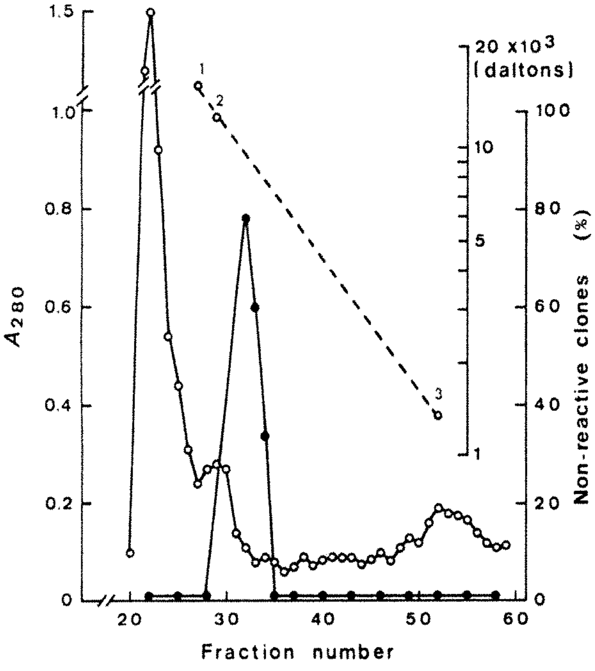


Fig. 1 *a*, Sephadex G-50 elution profile of the soluble fraction of sexually immature cells. A 2.5-ml sample of the supernatant obtained by homogenization and 105,000*g* centrifugation of immature cells (9×10^7) was applied to a Sephadex G-50 (Pharmacia, fine) column (2.5×70 cm) and chromatographed at 5 °C. Each fraction (4.4 ml) was assayed by injecting 15-pl samples into mature cells. ○, 280-nm absorbancy; ●, activity of immaturin (% of non-reactive clones brought about by the injection). Broken line, calibration curve of MWs: 1, myoglobin, 18,500; 2, cytochrome *c*, 13,000; 3, cyanocobalamin, 1,355 (all Sigma). *b*, DEAE elution profile. Active fractions (nos 31–33) from the Sephadex G-50 filtration were applied on a DEAE-Sephadex A-25 column (0.9×15 cm) containing 4 ml gel. Before application of the sample, the column was equilibrated with 50 ml solution A (see text). The elution medium was a linear gradient of from 0 to 500 mM NaCl in solution A. Size of each fraction was 4.4 ml ○, 280-nm absorbancy; ●, activity; broken line, concentration of NaCl.

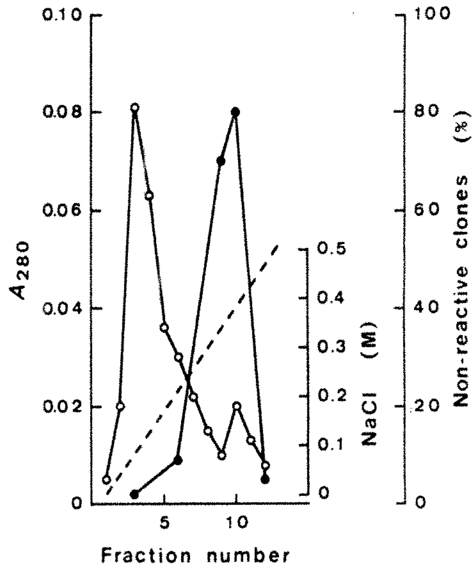


Fig. 2 Polyacylamide-gel pattern of the peak fraction of the DEAE chromatography (see text for details).

The MW of immaturin was estimated to be about 10,000 from Sephadex G-50 chromatography by comparing with marker molecules (cyanocobalamin, cytochrome *c* and myoglobin, Fig. 1). Determination of the MW using SDS-polyacrylamide gel electrophoresis was unsuccessful because the standard curve of MWs <11,000 was not linear. Immaturin lost its activity when exposed to polyacrylamide gel-bound insoluble trypsin (Sigma) for 65 min at 27–28 °C, whereas no loss of activity was observed when exposed to heat-inactivated, insoluble trypsin or agarose-bound, insoluble ribonuclease (Sigma) (Table 2). The activity was completely destroyed when exposed to 100 °C for 6 min.

Table 2 Activity of immaturin when treated with enzymes and heat

Treatment	No. of cells injected	No. of non-reactive clones	Activity (%)	Confidence interval (90%)
None	28	22	78.6	62–92
Trypsin	28	1	3.4	0.1–14
Heated trypsin	25	17	68.0	48–84
Ribonuclease	26	14	53.8	36–72
100 °C, 6 min	26	0	0	0–10.2
No injection	30	0	0	0–10

Insoluble trypsin (2.5 U, Sigma) or insoluble ribonuclease (5 U, Sigma) was added to fraction no. 9 from the DEAE-Sephadex column, in a solution containing 10 mM KCl, 5mM $MgCl_2$ and 10 mM Tris-HCl (*p*H 7.4), and incubated in a waterbath with shaking at 27–28 °C for 65 min. The particle-bound enzymes were removed by centrifugation at about 1,000*g* for 5 min and the supernatants were assayed for immaturin activity by injecting about 15 pl into mature cells. Heated trypsin was obtained by heating at 100 °C for 6 min. Activity is per cent of mature clones which contained more than 25% immature cells after injection (here called non-reactive clones).

From these results we conclude that immaturin is a heat-labile protein, which we isolated almost free of other proteins. We did not rigorously test the sample for purity from organic substances other than protein, but the purification procedure described above probably removed most of them. As repeated washing by ultrafiltration through a Diaflo membrane UM2 (Amico) at 4 °C under a pressure (2.0 kg cm^{-2}) of nitrogen gas did not decrease the activity of immaturin, it is unlikely that small molecules attached to protein have the immaturin activity. This experiment also seems to eliminate the possibility that loosely bound, small molecules are cofactors of immaturin, though it does not completely eliminate the possibility that tightly bound small molecules are cofactors or the site of immaturin activity.

Mating ability in *Paramecium* depends primarily on the presence of sexual cell-recognition molecules, located on the ciliary membranes and called the mating substances[12,13]. Specificity of the mating substances is known to be controlled genetically[9,14,15]. Immaturin evidently inhibits the expression of mating substance activity but the mechanism of its action, or its target of action within a cell, remains to be investigated. Immaturin may be comparable to several factors controlling gene expression, such as insulin and steroid hormone receptors in higher animals[16,17], repressors in *Escherichia coli*[18], CGA of *E. coli*[19], and the inhibitor of rRNA synthesis in *Xenopus*[20]. Isolation and characterization of immaturin should facilitate study of mechanisms of genic control of mating ability in *Paramecium*, of genic regulation of life-cycle phases in clonal ageing, and of cell differentiation in general.

We thank T. M. Sonneborn, D. L. Cronkite and G. H. Beale for helpful comments on the manuscript. This work was supported by grant 348001 from the Ministry of Education, Science and Culture to K.H.

Received 21 July; accepted 30 September 1980.

1. Sonneborn, T. M. in *The Species Problem* (ed. Mayr, E.) 155–324 (American Association for the Advancement of Science, Washington DC, 1957).
2. Miwa, I. & Hiwatashi, K. *Jap. J. Genet.* **45,** 269–275 (1980).
3. Bleyman, L. K. in *Developmental Aspects of the Cell Cycle* (eds Cameron, T. L., Padilla, G. M. & Zimmerman, A. Z.) 67–91 (Academic, New York, 1971).
4. Myohara, K. & Hiwatashi, K. *Genetics* **90,** 227–241 (1978).
5. Miwa, I., Haga, N. & Hiwatashi, K. *J. Cell Sci.* **19,** 369–378 (1975).
6. Miwa, I. *J. Cell Sci.* **36,** 253–260 (1979).
7. Miwa, I. *J. Cell Sci.* **38,** 193–199 (1979).
8. Miwa, I. *Zool. Mag., Tokyo* **88,** 622 (1979).
9. Hiwatashi, K. *Genetics* **58,** 373–386 (1968).
10. Sonneborn, T. M. *J. exp. Zool.* **113,** 87–147 (1950).
11. Davis, B. J. *Ann. N. Y. Acad. Sci.* **121,** 404–436 (1964).
12. Hiwatashi, K. in *Fertilization* Vol. 2 (eds Metz, C. B. & Monroy, A.) 255–293 (Academic, New York, 1969).
13. Kitamura, A. & Hiwatashi, K. *J. Cell Biol.* **69,** 736–740 (1976).
14. Sonneborn, T. M. in *Handbook of Genetics* Vol. 2 (ed. King, R. C.) 469–594 (Plenum, New York, 1974).
15. Siegel, R. W. & Larison, L. L. *Proc. natn. Acad. Sci. U.S.A.* **46,** 344–349 (1960).
16. Czech, M. P. *A. Rev. Biochem.* **46,** 359–384 (1977).
17. O'Malley, B. W. & Means, A. R. *Science* **183,** 610–620 (1974).
18. Gilbert, W. & Muller-Hill, B. in *The Lactose Operon* (eds Beckwith, J. & Zipser, D.) 93–109 (Cold Spring Harbor Laboratory, New York, 1970).
19. Riggs, A. D., Reines, G. & Zubay, G. *Proc. natn. Acad. Sci. U.S.A.* **68,** 1222–1225 (1971).
20. Siokawa, K. & Yamana, K. *Devl Biol.* **47,** 303–309 (1975).
21. Lowry, O. H. *et al. J. biol. Chem.* **193,** 265–275 (1951).

Lyt 1^+23^- cells appear in the thymus before Lyt 123^+ cells

Bonnie J. Mathieson*, Susan O. Sharrow†, Yvonne Rosenberg* & Ulrich Hämmerling‡

* Laboratory of Microbial Immunity, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20205
† Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205
‡ Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Most thymocytes are either immature or functionally deficient and express a series of lymphocyte cell-surface antigen markers designated Lyt 1, Lyt 2 and Lyt 3 (refs 1, 2) which have been useful in distinguishing functional subsets of T cells[3,4]. In contrast, a small population of cortisone-resistant thymocytes (CRT), confined to the thymic medulla after acute corticosteroid treatment are functionally more mature. These cells, like peripheral T cells, have restricted expression of Lyt antigens and mostly are either Lyt 1 or Lyt 123 cells[3–5]. It has thus been assumed that all thymocytes initially are Lyt 1^+, 2^+, 3^+ and by differentiation lose either Lyt 1 or Lyt 2, 3 to result in Lyt $1^+(23^-)$ and Lyt $(1^-)23^+$ cells[6–8]. Using immunofluorescence (IF) and flow microfluorometry (FMF) analyses to detect Lyt antigen expression quantitatively without the requirement for cell lysis, we have now re-examined the expression of Lyt 1, 2 and 3 antigens on murine fetal thymocytes from 14 to 19 days of gestation and on normal thymocytes from birth to 2–3-month-old adults. These studies demonstrate that cells with the Lyt 1^+23^- phenotype first appear in the thymus several days before Lyt 123^+ thymocytes are detected, and suggest either a micro-environmental or site-specific influence for phenotypic differentiation and/or two independent, pre-committed lineages.

The details of the methods used for FMF detection of IF for Lyt have been reported previously[5]. For these studies, in addition to alloantisera prepared in our laboratory, monoclonal reagents for Lyt 1.1 and Lyt 2.2 specificities have been used to avoid any problems of nonspecific reactions. In our experience, these monoclonal reagents give patterns of reactivity both in adult lymphoid tissue distribution and strain distribution equivalent to the related, allelically specific antisera described by others[9].

Data from one experiment are shown in Table 1. With fetal thymocytes from C57BL/6-Ly-1^a mice, an increased proportion of Lyt 1^+ thymocytes are observed between 15 and 19 days of gestation. Using monoclonal anti-Thy 1.2, 80% of the fetal thymocytes are positive for Thy 1 expression at the 15 and 16 day points, whereas >97% are positive for Thy 1 at later time points (data not shown). These positive cells also express quantitatively increased amounts of Lyt 1 antigen. This latter conclusion is drawn from the increase in median fluorescence units (FU) for the thymocytes at each time point and can be seen more explicitly in data from an additional experiment in Fig. 1. This increase in the level of Lyt 1 expression per cell occurs despite the fact that cell size is decreasing through this period, as reflected by the decrease in the median light scatter units. For comparison, these same parameters are shown for adult thymocytes. Thus, in contrast to adult thymocytes[10], the quantitative level of the Lyt 1 phenotype is not related to cell size.

Until day 17 of gestation, almost no cells expressing Lyt 2 or Lyt 3 antigen are seen (Table 1). Figure 2 shows the IF profiles of Lyt 3.2 expression on thymocytes from embryos at days 16 and 17 of gestation. In a series of six experiments, a marked increase in the proposition of thymocytes expressing both Lyt 2 and Lyt 3 is seen (7–15% above control background on days 14–16 to >70% on day 17). The adult phenotype, 10% Lyt 1^+23^- : 90% Lyt 123^+ which we previously described[5], is observed at 18 days, about 24 h before birth in this strain. At no time are all the fetal thymocytes Lyt 123^+.

The dramatic appearance of the Lyt 123^+ cells on day 17 of gestation follows differentiation of cortical histological sites in the murine thymus observed by others at 15 days of gestation[11,12]. The appearance of these cells also corresponds to a marked increase in cell number which occurs at this time. These data suggest either than Lyt 123^+ cells are derived from Lyt 1^+23^- precursors, or that two separate pre-committed lineages for Lyt 1^+23^- and Lyt 123^+ cells in the thymus take up residence at different times during ontogeny. The concept of two independent lineages of cells, either or both of which may

0028–0836/81/030179—03$01.00

Table 1 Lyt expression on fetal thymocytes

Fetal age (days)	Lyt 1 Minimum % positive	Lyt 1 Median FU of positive cells	Lyt 2 % positive	Lyt 3 % positive	Median LSU of viable cells	% Large cells	No. of cells per thymus
15	50.1	201 (77)	6	ND	606	70	4×10^5
16	51.8	236 (69)	15	13	600	68	5.7×10^5
17	95.8	375 (42)	72	73	564	63	3.3×10^6
19	96.5	365 (39)	79	80	516	52	4×10^6
Adult	98.4	439 (31)	91	91	440	21	$1\text{-}2\times10^8$

Fetal age was determined by calculating from the date of detection of a vaginal plug (plug date = day 0). Minimum % positive was determined by the number of stained cells above the intersection of the IF profile of anti-Lyt 1.1 plus fluoresceinated goat anti-mouse immunoglobulin (FlGAMIg) stained cells and the IF profile of the cells treated with FlGAMIg alone. This type of intersection is seen in Fig. 1. The median values for Lyt 1^+ cells are calculated above the intersection (145 FU) for this experiment. FU are calculated as previously described[10]. For comparison the value for surface immunoglobulin on B cells is ~6,000 FU. Median FU for the controls are given in parentheses next to the FU values for Lyt 1^+ cells. Light scatter units (LSU) are arbitrary units of measurement reflecting cell size which have been described by others and which we have used previously for comparison of thymocytes[5,27]. Large thymocytes are reduced to about 20% of the total thymocyte population within 1 week of birth. The number of cells per fetal thymus was determined from the total cell recovery of 7–12 embryos at each time point. ND, not determined.

subsequently be subdivided, is certainly not a new proposal for thymocyte differentiation. However, the characterization of these two lineages with different thymocyte phenotypes in the fetal development of the thymus has not been demonstrated previously.

This finding that Lyt 1^+23^- cells are present in the fetal thymus before Lyt 123^+ cells contrasts with the previously proposed concept that all thymocytes are Lyt 123^+ before differentiative loss of Lyt 1 or Lyt 2 and Lyt 3 (ref. 4). After birth, CRT[3–5], the thymus migrant population[13] and thymocytes non-agglutinated with peanut lectin[10] all contain both Lyt 1^+23^- and Lyt 123^+ cells, presumably in a more mature state of differentiation. In normal adult thymocytes, Lyt 1^-23^+ cells are not seen by IF with the quantitative methods used here[5,10]. Therefore, if one still wishes to assume that all thymocytes and T cells differentiate through an Lyt 123^+ state, it is required that at least during ontogeny Lyt 1^+23^- cells generate Ly 123^+ cells, and from those, more mature Lyt 1^+23^- cells and Lyt 123^+ cells which exit from the thymus. These data do not exclude the possibility that some Lyt null precursors express all three Lyt antigens simultaneously, nor do they exclude the possibility that a small number of Lyt 123^+ precursors are instantly driven to Lyt 1^+23^- in early ontogeny.

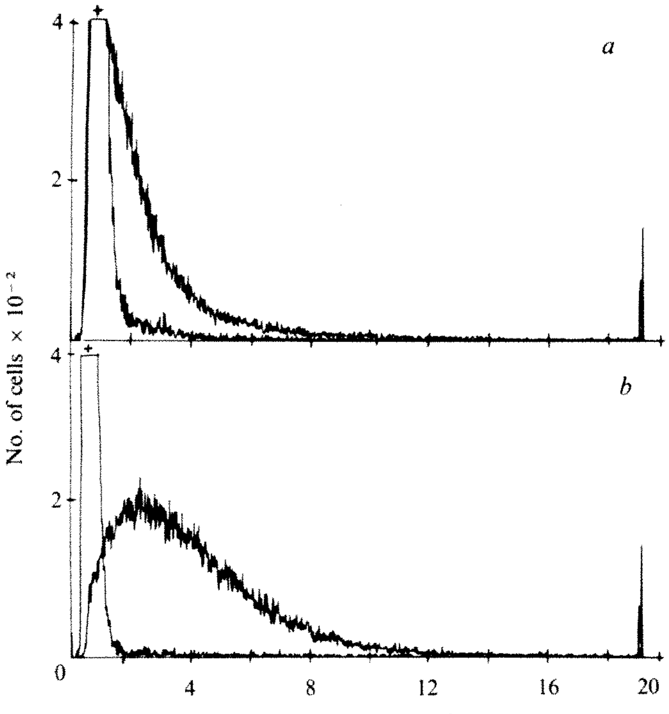


Fig. 1 Fetal thymocytes from B6-Ly 1^a mice which express the Lyt 1.1 antigen were treated with anti-Lyt 1.1 monoclonal antibody (clone line 39/107) and FlGAMIg. The details of the fluorescent reagents and the IF procedure are published elsewhere[5] as are the details of the use of the FACS-II (Becton Dickinson) for quantitative FMF analysis[5,27]. The data configurations presented in *a* and *b* as the IF profiles on 50,000 cells per sample are also described elsewhere[28]. In *a*, 15-day fetal thymocytes, the positive cells (cells with >160 FU), are compared with negative control fetal thymocytes from B6 mice which express the Lyt 1.2 antigen. As seen in *b*, the fetal thymocytes at day 17 of gestation are much more positive, with a reduced overlap between the positive cells and the negative control, in this case FlGAMIg treatment alone.

The fetal thymus of the mouse develops in a particular histological pattern of differentiation. Initially, the thymic rudiment appears as a largely epithelial structure with a sparse lymphoid population, comparable in both lymphoid type and size to the medullary area of the differentiated adult thymus. Subsequently, the densely lymphoid cortical areas which are largely composed of small thymocytes appear and a particular histological type of cortical epithelial cell develops in the fetal thymus[11,12,14]. It is almost certain that some stem cells in the early fetal thymus are capable of differentiating into Lyt 123^+ cells, because Owen and Ritter[14], Mandel and Russell[11,12], and Kamarck and Gottlieb[8] have observed small cortical thymocyte differentiation in murine fetal thymus organ cultures deprived of traffic of additional cells from the periphery. Furthermore, Kamarck and Gottlieb[8] indicated that Lyt 123^+ cells could differentiate in these fetal organ cultures. This is being tested further with fetal thymus transplants, which are derived from mice expressing allelically different Lyt specificities.

Thus, the microenvironment of the thymus may be critical for phenotype generation. When the early 'medullary' epithelial environment is predominant, perhaps only Lyt 1^+23^- cells can develop. Once the cortical epithelial cells[11,12,14] have differentiated an appropriate site for Lyt 2, 3 expression, these antigens may then be detected on populations subsequently differentiating in the thymus. This would presume a common precursor with no site-specific homing, but site-specific differentiation would predict that adult thymic repopulation studies would show the adult Lyt phenotypes and not the sequential pattern seen in ontogeny.

The alternative possibility of two separate lineages would more easily explain both the finding that there are, from an early time point in ontogeny, Lyt 1^+23^- thymocytes as demonstrated here, and that a small proportion of thymocytes with this phenotype are found throughout adult life[5]. This could also explain other phenotypic subset analyses of the differential expression of Thy 1 in the thymus[15], as well as sequential reactivity of fetal thymocytes to phytohaemagglutinin and concanavalin A[16].

Avian thymus studies indicate that the differentiation involves two lineages which may even have different pre-thymic path-

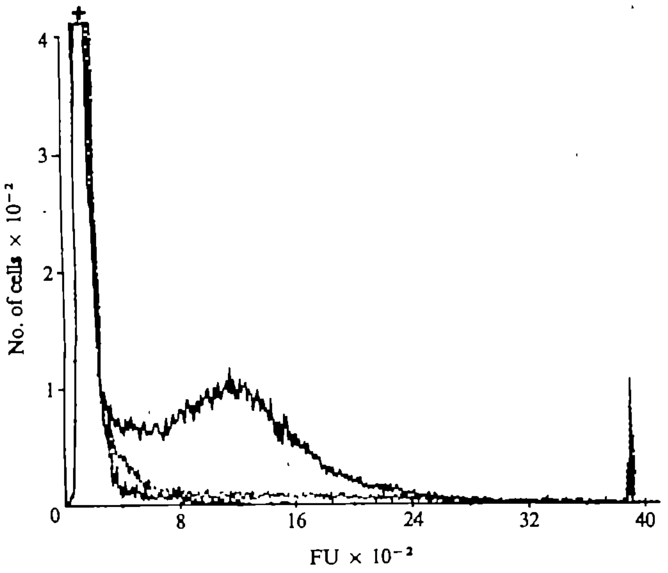


Fig. 2 Fetal thymocyte from B6-Ly-1^a mice which express the Lyt 3.2 antigen were treated with anti-Lyt 3.2 serum[9] and FlGAMIg antibodies. At 17 days gestation (——), the IF is seen as a peak of positive cells whereas at day 16 (----) the IF is less than 10% above the 16-day fetal thymocytes treated with FlGAMIg (—·—·) alone. The median value of the Lyt 3^+ cells was not noticeably changed in these experiments and varied only between 1,200 and 1,600 FU. The FlGAMIg control has 3.8% cells >400 FU and 1 6% of cells >800 FU.

ways[17,18]. An early cohort of thymic cells passes through the bursa of Fabricius, a B-cell differentiation site in birds, and takes up residence in the thymus. Cells, which populate thymus at a later time in development, apparently bypass the bursa. This parallels the two-stage distinct sequential appearance seen for the two Lyt phenotypes reported here. One might even consider that precursors of Lyt 1^+23^- cells pass some similar B-cell site at which a T-helper function is in some way induced or programmed in the lineage. This is clearly hypothetical in the mouse, because no comparable unified B-cell generation site is known. However, there is evidence of separated functions of isolated subpopulations of Lyt differentiated peripheral T cells[19].

Finally, a combination model is proposed in which Lyt 1^+23^- cells give rise to two separate lineages only one of which 'seeks' a cortical site or pathway. This Lyt 123^+ lineage may require multiple divisions in the cortex for both functional and phenotypic differentiation and for repertoire development. The requirement for cell division has been documentated in other differentiation systems[20]. This model of separate lines with site-specific differentiation would be consistent with data and models of thymocyte differentiation presented by several investigators, including Weissman[21], Shortman and Jackson[15], Zinkernagel[22] and Stutman[23]. Although Leckband and Boyse[24] previously indicated a correlation with TL^- phenotype of CRT, a presumably more mature medullary thymic subpopulation, recent observations of Kisielow *et al.*[25] show TL^+ cells with differentiated Lyt phenotypes. Data obtained from analysis of thymic lymphomas[26] led us to propose earlier that such immature thymocytes might exist which were TL^+ but differentiated for Lyt phenotype. Thus, we favour this latter combination model which presumes independent pre-committed lineages with site-specific differentiation and therefore presumes nothing about the maturity of cells found in either the cortical or medullary site in the thymus.

We thank the members of the Laboratory of Microbial Immunity who have encouraged and supported our work, especially Dr R. Asofsky and Ms B. J. Fowlkes for provision of fluorescent reagents used in these studies, and V. Shaw and O. Childers for preparation of the manuscript. We also thank Dr E. A. Boyse for provision of the original breeding stocks of the B6-Ly congeneic mice, Dr M. Potter for provision of experimental mice, through a NCI contract, and Drs J. Wunderlich, N. Jerne and H. Holtzer for stimulating discussions during this work.

Received 18 July; accepted 4 November 1980

1 Boyse, E A , Miyazawa, M , Aoki, T & Old, L. J. *Proc R. Soc* B**170**, 175–193 (1968)
2 Boyse, E A , Itakura, K , Stockert, E , Iritani, C A & Miura, M *Transplantation* **11**, 351–353 (1971)
3 Kisielow, P *et al Nature* **253**, 219–220 (1975)
4. Cantor, H & Boyse, E A *J exp Med* **141**, 1376–1389 (1975)
5 Mathieson, B J , Sharrow, S O , Campbell, P S & Asofsky, R *Nature* **277**, 478–480 (1979)
6. Cantor, H & Boyse, E A *Immun Rev* **33**, 105–124 (1977)
7 Shen, F W , McDougal, J S , Bard, J & Cort, S P *J. exp Med* **151**, 566–572 (1980)
8. Kamarck, M E & Gottlieb, P D. *J Immun* **119**, 407–415 (1977)
9 Shen, F -W , Boyse, E A & Cantor, H *Immunogenetics* **2**, 591–595 (1975)
10 Betel, I , Mathieson, B J , Sharrow, S O & Asofsky, R *J Immun* **124**, 2209–2217 (1980)
11 Mandel, T *Z. Zellforsch* **106**, 498–515 (1970)
12. Mandel, T & Russell, P J *Immunology* **21**, 659–674 (1971)
13 Scollay, R , Kochen, J , Butcher, E & Weissman, I *Nature* **276**, 79–80 (1978)
14 Owen, J J T & Ritter, M A *J exp Med* **129**, 431–437 (1969)
15. Shortman, K & Jackson, H *Cell Immun* **12**, 230–246 (1974)
16 Mosier, D E *J Immun* **112**, 305–310 (1974)
17. LeDouarin, N M & Jotereau, F V *J exp Med* **142**, 17–40 (1975)
18 Jotereau, F V & Houssaint, E in *Developmental Immunobiology* (eds Solomon, J. B & Horton, J D) 123–130 (Biomedical Press, Amsterdam, 1977)
19 Huber, B , Cantor, H , Shen, F -W & Boyse, E A *J exp Med* **144**, 1128–1133 (1976)
20 Holtzer, H. in *Stem Cells and Tissue Homeostasis* (eds Lord, B I , Potten, C S & Cole, R J) 1–27 (Cambridge University Press, 1978)
21 Weissman, I L *J exp Med* **126**, 291–304 (1967)
22 Zinkernagel, R. *Immun Rev* **42**, 224–270 (1978)
23 Stutman, O *Immun Rev* **42**, 138–184 (1978)
24 Leckband, E & Boyse, E A *Science* **172**, 1258–1260 (1971)
25. Kisielow, P , Draber, P & Wysocka, M *Eur J Immun* **9**, 1023–1025 (1979)
26 Mathieson, B J , Campbell, P S , Potter, M & Asofsky, R *J exp Med* **147**, 1267–1279 (1978)
27 Loken, M R & Herzenberg, L A *Ann N Y Acad Sci* **254**, 163–171 (1975)
28 Miller, M H , Powell, J I , Sharrow, S O & Schultz, A R *Rev scient Instrum* **49**, 1137–1142 (1978)

Selective inhibition of T suppressor-cell function by a monosaccharide

Ulrich H. Koszinowski* & Michael Kramer

Institut für Immunologie und Genetik, DKFZ, 6900 Heidelberg, FRG

Interactions between regulatory T lymphocytes and other cells are assumed to occur at the level of the cell surface. T cells which suppress the generation of specifically effector cells have been described as having antigenic, idiotypic, allotypic and I-region specificity[1–4]. Other T suppressor cells generated by *in vitro* cultivation with or without mitogenic stimulation[5,6] have suppressive activity for T and B cells but no specificity can be assigned to them. These T suppressor cells (T_s) inhibit various lymphoid functions—this either reflects their polyclonal origin or indicates that the structures recognized by the T_s receptors must be common for many cell types. Carbohydrates on cell membrane-inserted glycoproteins or glycolipids might function as specific ligands for recognition by cellular receptors or soluble factors. Almost all cell-surface proteins of mammalian cells are glycosylated. There is evidence for lectin-like carbohydrate binding proteins not only in plants[7] but also in toxins[8], viruses[9], prokaryotic cells[10] and even mammalian cells, including T cells[11]. A functional role for these lectin-like proteins has been described for slime moulds and suggested for the selective association of embryonic cells[12,13]. We report here that addition of a monosaccharide can counteract the effect of T suppressor cells during the generation of alloreactive cytotoxic T cells (CTLs) *in vitro*.

---

* Present address. *Federal Research Institute for Animal Virus Diseases*, 7400 Tubingen, PO Box 1149, FRG

0028-0836/81/030181—04$01 00

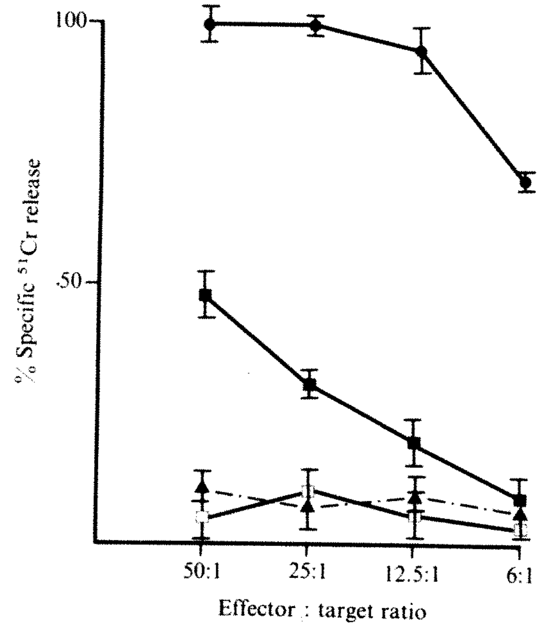


Fig. 1 Regulatory effects of MM on suppressor-cell function. Alloreactive CTLs were generated from C57BL/6J responder spleen cells by incubation with 2,000 rad-irradiated allogeneic DBA/2J spleen cells (6–8 week-old mice from the Institut für Versuchstierzucht, Hannover). Cells were cultivated in 24 flat-bottom Linbro plates (Flow) at a responder/alloantigen ratio of 1 : 1 and a concentration of 2.5×10^6 responder cells per ml RPMI medium supplemented with 10% heat-inactivated fetal calf serum and antibiotics. Cells were collected 6 days later and tested for cytolytic activity in the ^{51}Cr-release assay[17] using P-815 mastocytoma cells as targets. ●, Suppressor cells (T_s) from C57BL/6J mice were generated by precultivation in the absence of alloantigens using the method of Rollwagen and Stutman[15] with the modification that erythrocyte lysis was omitted. When tested for cytolytic activity after 6 days, these cells were found to be negative (▲). After addition of these precultivated cells to new cultures containing responder cells and alloantigens at the ratio of responder/T_s varying from 1 : 2 to 1 : 8 (this experiment 1 : 4), production of alloreactive CTLs was abolished (□). Addition of 10 mg ml^{-1} MM (Roth; Sigma; Calbiochem; Hoechst) counteracted the suppressive effect (■). Experiments were carried out in triplicate. Bars indicate standard deviations.

Nonspecific suppressor-cell production in *in vitro* cultivated lymphocytes by the plant lectin concanavalin A (Con A) can be inhibited by the addition of methyl-α-D-mannopyranoside (MM) during the induction phase[14]. We wondered whether similar phenomena could be demonstrated in a system without Con A. T_s were produced *in vitro* in cultures of spleen lymphocytes incubated in complete medium without antigenic stimulation for 6 days, as described elsewhere[15]. Suppressor cells produced in this way inhibit a variety of immune responses nonspecifically and without H-2 restriction. When added in graded numbers to a primary mixed leukocyte culture (MLC) on day 0, these cells were able to suppress or abolish the generation of alloreactive CTLs. This inhibitory activity was neither due to cell crowding effects (data not shown) nor to cytolytic activity in the suppressor cell populations[16]. Data from a representative experiment showed that generation of CTLs is abolished in the presence of suppressor cells (Fig. 1) and that these cells were not cytolytic after cultivation with antigen. Addition of 10 mg per ml MM to the culture system counteracted the suppressive effect, resulting in the reconstitution of a cytolytic response.

Treatment of the transferred cells with anti-Thy-1.2 and complement abolished their suppressive effects, demonstrating that the MM-sensitive suppression is a T cell-dependent phenomenon (Fig. 2).

Table 1 Ability of methyl-α-D-mannopyranoside (MM) to inhibit the function of suppressor cells

Molar concentration of sugar during CTL induction	Ratio of responder: suppressor cells	Ratio of killer: target cells required for 1 LU	r^2	% T_s activity
—	1 : 0	10.0	0.97	0
—	2 : 1	39.1	0.95	100
10^{-2}	2 : 1	15.8	0.99	20
10^{-3}	2 : 1	17.2	0.98	25
10^{-4}	2 : 1	18.6	1.00	30
10^{-5}	2 : 1	27.2	0.97	59

Methyl-α-D-mannopyranoside (Roth) was re-crystallized and added at different molar concentrations to the culture containing responder cells, allogeneic stimulator cells and suppressor cells. Effector cells produced in these conditions were tested for cytolytic activity at 4 effector: target cell ratios (third column). The killer: target cell ratio required for 1 lytic unit (LU), arbitrarily defined as the ratio resulting in 33% specific lysis in the 4-h assay period, was calculated by logarithmic curve fitting[17]. *r*, Coefficient of determination. The cytolytic capacity of suppressed responder-cell cultures was taken as 100% T_s activity.

Addition of the carbohydrate alone to the culture medium did not inhibit the generation of alloreactive CTLs. MM therefore does not interfere with antigen recognition, differentiation and expansion of alloantigen-specific T cells (Fig. 3*a*). Furthermore, no inhibition by MM was seen in the effector phase of cytolytic T

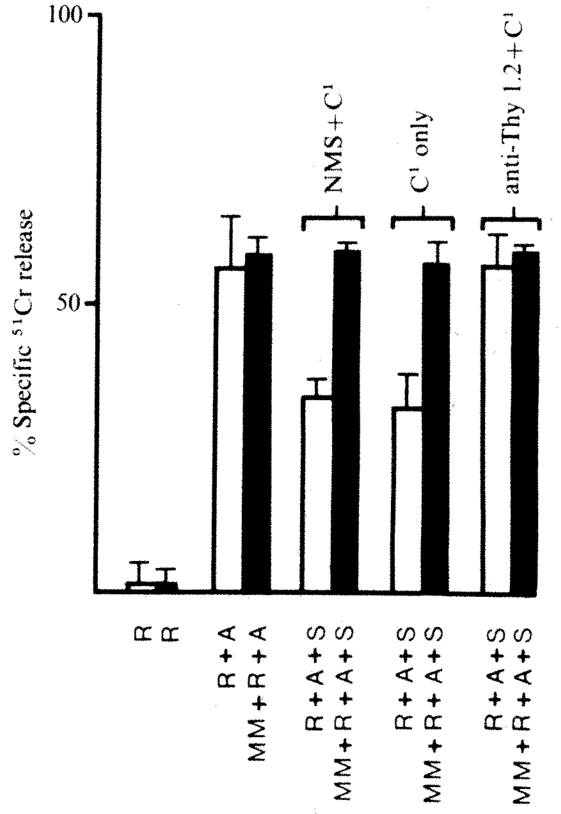


Fig. 2 Characterization of the MM-susceptible suppressor cells. Generation of alloreactive CTLs in presence or absence of MM and suppressor cells. After the 6-days induction phase, suppressor cells were incubated with either normal mouse serum (NMS) or anti-Thy-1.2 serum (1 : 20 final dilution) for 30 min at 20 °C, centrifuged and resuspended in freshly thawed selected rabbit serum as a source of complement. Further incubation was at 37 °C for 30 min. R, C57BL/6J responder cells (2.5×10^6 ml^{-1}); A, DBA/2J allogeneic cells (2.5×10^6 ml^{-1}); S, suppressor cells (0.6×10^6 ml^{-1}); MM, methyl-α-D-mannopyranoside (10 mg ml^{-1}). Effector: target cell ratio, 25 : 1; target, P-815 mastocytoma cells.

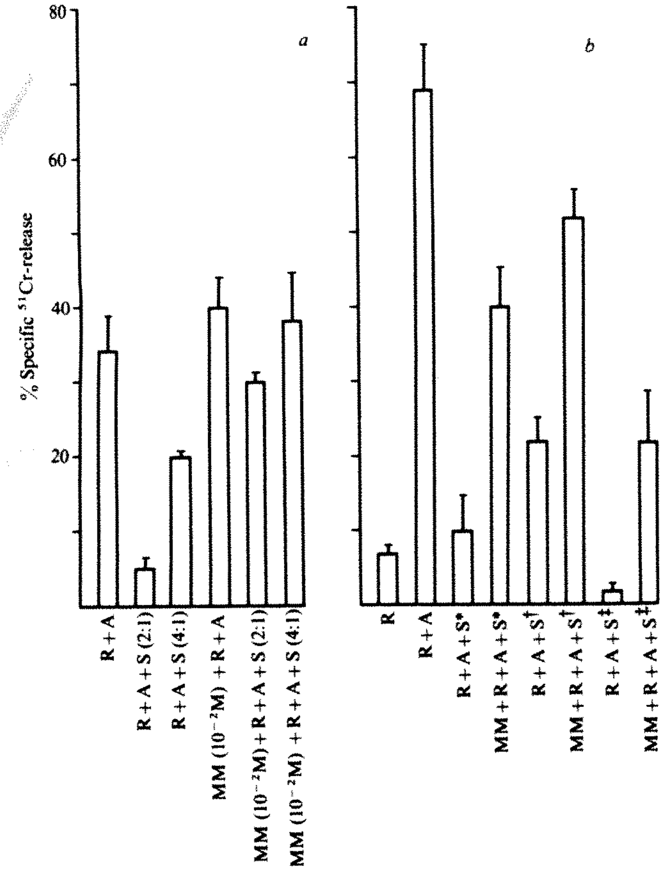


Fig. 3 Effect of MM on the generation of effector T cells. *a*, Effect on killer-cell generation. Responder cells (R) in presence of allogeneic cells (A) (see Fig. 1 legend) generate CTLs (first column). This effector-cell generation is suppressed in presence of suppressor cells (S) added at different R:S ratios at the start of culture. Addition of MM (10^{-2} M) inhibits suppressor-cell function (columns 5, 6) but not the generation of CTLs in the absence of suppressor-cells (column 4). *b*, Effect on suppressor-cell generation. * Suppressor cells (S) were generated as described in Fig. 1 legend and added to the culture to suppress CTL generation in presence or absence of 10 mg per ml MM. † Suppressor cells were cultivated for the first 72 h of the 6-day induction period in the presence of MM (10 mg ml^{-1}) and then tested for function in presence or absence of MM. ‡ Suppressor-cell generation in presence of MM during the 6-day induction period. Effector/target cell ratio in both experiments, 50:1; responder/suppressor cell ratio in *b*, 4:1.

cells because target cell lysis was the same in the presence or absence of MM (10 mg ml^{-1}) (data not shown). Similarly MM did not affect the induction phase of suppressor-cell generation (Fig. 3*b*). Spleen cells were cultivated in presence of MM either for the initial 72 h or the whole 6-day induction period before addition to the second culture. Both populations were found to have comparable suppressive activity when transferred to the MLC. This indicates that MM does not affect the generation of T_s from their precursors. Furthermore, we found no evidence for reduced DNA synthesis (^{3}H-thymidine incorporation) or decreased viability of T_s in culture in the presence of MM (data not shown). However, the capacity of the suppressor cells to inhibit a primary CTL was inhibited when MM was added to the second culture (Fig. 3*b*). Due to the different susceptibility to carbohydrate in the culture medium, our results clearly define cytotoxicity and suppression in this system as two separable manifestations of effector T-cell functions, a fact which is only shown with difficulty during regulation of CTL generation by T_s[18].

The effect of MM on T_s effector function could be attributed to the carbohydrate and not to impurities of the reagent used because MM from different sources (data not shown) and also the re-crystallized sugar were equally effective (Table 1). Osmotic effects seem unlikely because several other monosaccharides were ineffective (Fig. 4). A toxic effect of MM on lymphocytes is also unlikely because concentrations even 100-fold higher than those active in our system were not toxic to leukocyte function[19].

Our study demonstrates that cell interactions can be blocked by monosaccharides. Only the regulatory function of T_s is counteracted—neither CTL effector function nor the generation of T_s and CTLs is inhibited. At this stage it is impossible to determine how the T_s interact with their targets during CTL activation. Alternative mechanisms must be considered. The sugars may sterically interfere with structures involved in recognition of target cells by T_s. MM could also have reversible selective inhibitory effects on T_s metabolism, affecting function but not differentiation, perhaps by interfering with the glycosylation of recognition structures. We favour the hypothesis that recognition structures on T_s or secreted T_s products, by interaction with complementary structures on their target cells, mediate suppression and that the carbohydrate competitively blocks this interaction. Results obtained with stem cells[20] and T cells[21] suggest that differentiation may be accompanied by a change in carbohydrate binding capacity of cell-surface components.

The fact that D-mannose is often found in mammalian glycoproteins either as core or end groups[22] does not mean that the specificity resides only in this sugar. Other mono-, and in particular, oligosaccharides should be tested for specificity by determining the concentration required to inhibit the regulative T-cell interaction. The advantage of our system is that it studies functional interactions. However, one disadvantage is that the incorporation of sugar into cells may decrease their capacity to compete at the extracellular level and the concentration of

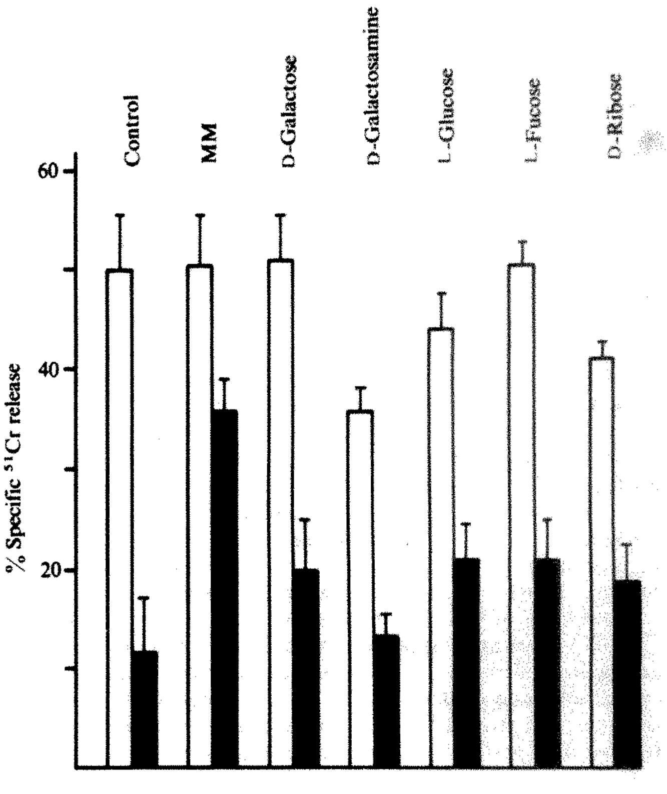


Fig. 4 Specificity of the carbohydrate effect on suppressor-cell functions. For suppressor and responder culture conditions, see Fig. 1 legend. The control shows the activity of responder cells incubated in absence (open columns) or presence (solid columns) of suppressor cells. In the other experimental groups different carbohydrates were added at 10^{-3} M concentration during the induction phase to the suppressed and non-suppressed cell cultures.

carbohydrates not toxic to effector-cell generation may not be sufficient to inhibit measurably suppression.

No definite function has been ascribed to oligosaccharides although their role in cell recognition has been suggested. Our observations on the counteraction of suppressor-cell activity suggest that carbohydrates play a part in T-cell regulation.

We thank Ms Ulrike Ette for technical help, Ms Ursula Rohmann for preparation of the manuscript, and our colleagues, in particular Dr M. Simon, for their criticisms. This work was partially supported by DFG grants Ko 571/5 and 571/7.

Received 7 July; accepted 3 November 1980.

1. Benacerraf, B., Kapp, J. A., Debré, P., Pierce, C. W. & de la Croix, F. *Transplantn Rev.* **26**, 21–38 (1975).
2. Eichmann, K. *Eur. J. Immun.* **5**, 511–517 (1975).
3. Herzenberg, L. A., Okamura, K. & Metzler, C. M. *Transplantn Rev.* **27**, 57–83 (1975).
4. Tada, T., Taniguchi, M. & Takemori, C. S. *J. exp. Med.* **144**, 713–725 (1976).
5. Nadler, L. M. & Hodes R. J. *J. Immun.* **118**, 1886–1895 (1977).
6. Peavy, D. L. & Pierce, C. W. *J. exp. Med.* **140**, 356–369 (1974).
7. Lis, H. & Sharon, N. *A. Rev. Biochem.* **42**, 541–575 (1973).
8. Moss, J., Fishman, P. H., Manganiello, V. C., Vaughan, M. & Brady, R. O. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1034–1037 (1976).
9. Marchesi, V. T. & Andrews, E. P. *Science* **174**, 1247–1248 (1971).
10. Barondes, S. H. & Rosen, S. D. in *Neuronal Recognition* (ed. Barondes, S.) 331–356 (Plenum, New York, 1976).
11. Boldt, D. H. & Armstrong, J. P. *J. clin. Invest.* **57**, 1068–1078 (1976).
12. Geltosky, J. E., Wesemann, J., Bahke, A. & Lerner, R. A. *Cell* **18**, 381–398 (1979).
13. Moscona, A. A. in *Neuronal Recognition* (ed. Barondes, S.) 205–226 (Plenum, New York, 1976).
14. Dutton, R. W. *Transplantn. Rev.* **26**, 39–55 (1975).
15. Rollwagen, F. M. & Stutman, O. *J. exp. Med.* **150**, 1359–1366 (1979).
16. Ferguson, R. M., Anderson, S. M. & Simmons, R. L. *Transplantation* **26**, 331–339 (1978).
17. Koszinowski, U. & Simon, M. M. *Eur. J. Immun.* **9**, 715–722 (1979).
18. Orosz, C. G. & Bach, F. H. *J. Immun.* **124**, 1419–1421 (1979).
19. Rocklin, R. E. *J. Immun.* **116**, 816–820 (1976).
20. Grabel, L. B., Rosen, S. D. & Martin, G. R. *Cell* **17**, 477–484 (1979).
21. Lehmann, V., Streck, H., Minner, I., Krammer, P. H. & Ruschmann, E. *Eur. J. Immun.* (in the press).
22. McKenzie, I. F. C., Clarke, A. & Parish, C. R. *J. exp. Med.* **145**, 1039–1053 (1977).

Is erythropoietin the only factor which regulates late erythroid differentiation?

Barbara Fagg[*]

Biologie Cellulaire, Swiss Institute for Experimental Cancer
Research, CH-1066 Epalinges s./Lausanne, Switzerland

Haematopoiesis is a useful model system for studying differentiation and the regulation of precursor cell populations intermediate between the multipotential stem cell and terminally differentiated end cells. For many years, erythropoietin (Epo) was recognized as the hormone which controls red cell production *in vivo*[1]. Although other substances are now known to be required during the initial stages of erythropoiesis[2,3], late erythroid differentiation is regarded as strictly Epo-dependent[4]. This concept is supported by the recent demonstration that the addition of Epo alone to serum-free bone marrow cell cultures is sufficient to stimulate the CFU-E (colony-forming unit-erythroid—a late erythroid precursor cell approximating to a proerythroblast) to complete differentiation into mature erythrocytes.[5] However, the data reported here indicate that mature erythroid colonies (indistinguishable from those formed by CFU-E + Epo) are formed when adult bone marrow cells are grown for 2 d in methyl cellulose cultures containing spleen cell-conditioned medium (SCM) but no added Epo. SCM is a rich source of growth factors[6,7] and initial observations[8] suggested that its 'Epo-like' activity could be attributed to: (1) Epo, or (2) a factor which enhances the activity of small amounts of Epo in the culture medium, or (3) a factor(s) distinct from Epo which is also capable of stimulating late erythroid differentiation. The experiments reported here, which include a partial characterization of the 'Epo-like' activity, support the latter interpretation.

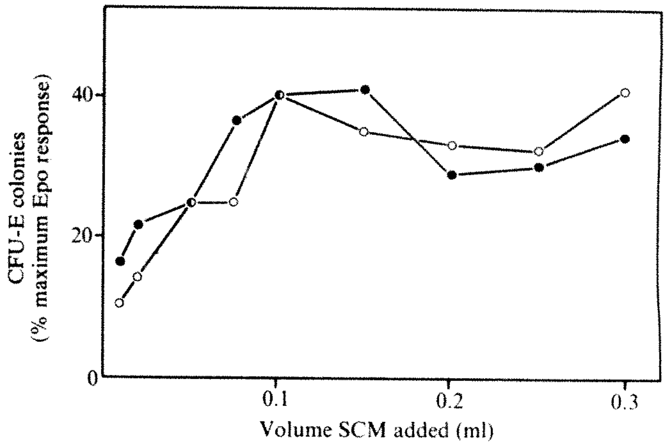


Fig. 1 Formation of erythroid colonies from adult mouse bone marrow cells grown for 2 d *in vitro* in the presence of different volumes of SCM. The number of CFU-E colonies is expressed as a percentage of the maximum response to Epo (0.15 units per ml culture; 307 ± 31 colonies per 10^5 cells, mean ± s.e.m. of 19 observations) in the same experiment. Epo was prepared from the urine of an aplastic anaemia patient according to Iscove *et al.*[5,12]. Either 20% (●) or 0% (○) fetal calf serum was added to the culture medium. Bone-marrow cell cultures contained 0.8% methyl cellulose, fetal calf serum (0% or 20%), 32 μg ml^{-1} lecithin, 18 μg ml^{-1} cholesterol, 3.6×10^{-6} M human transferrin, 1% deionized BSA and 8×10^4 cells ml^{-1}. They were evaluated as described by Iscove *et al.*[12].

SCM was collected from mouse spleen cells which had been stimulated with allogeneic irradiated spleen cells[9]. Adult DBA/2J mouse bone marrow cells grown for 7–10 d *in vitro* in the presence of SCM produced granulocyte/macrophage, megakaryocyte, erythroid (bursts) and mixed colonies. Similar results have been described previously using this type of SCM or that obtained from pokeweed mitogen(PMW)- or concanavalin A (Con A)-stimulated spleen cells[10,11]. The novel finding in the present study was that after 2 days *in vitro*, in the presence of SCM but no exogenous Epo, adult bone marrow cells produced colonies of 8–60 haemoglobinized (benzidine-positive) erythroid cells which were indistinguishable from those stimulated by Epo. Experiments designed to characterize this response in more detail showed that the number of erythroid colonies (cultured in the presence of 20% serum) increased in direct proportion to the amount of SCM added to the culture and reached a plateau value that was maintained at high SCM doses (Fig. 1). At a dose of SCM which maximally stimulated colony formation there was a linear relationship between the number of bone marrow cells plated ($0.25–4.0 \times 10^5$ cells per plate) and the number of erythroid colonies (correlation coefficient, $r = 0.98$). This indicates that the response occurs at the level of the individual cell and that colony growth is not dependent on cooperation between cells in the culture.

Fetal liver cells are often used to assay growth factors because they are generally more responsive and contain a higher proportion of precursor cells than adult tissues. SCM had a similar 'Epo-like' effect on 12-day fetal liver cells in culture to that on adult bone marrow cells. SCM from Con A- and PWM-stimulated spleen cells (one sample of the latter was provided by Dr G. R. Johnson, Melbourne, Australia) were also examined and found to contain a low level of 'Epo-like' activity when tested on adult bone marrow cells (data not shown).

All seven SCM preparations tested were active, although the level of 'Epo-like' activity varied between batches; the maximum response to SCM was between 41% and 80% of the maximum response to Epo, and required 33–150 μl of SCM per ml culture medium. The activity did not seem to be derived from the serum used to prepare SCM, because batches of SCM containing 1, 2 or 20% (v/v) serum gave similar results and

---

[*] Present address: Department of Microbiology, Medical Sciences I, University of California, Irvine, California 92717.

0028-0836/81/030184—03$01.00

equivalent amounts of the same serum tested alone did not exhibit any 'Epo-like' activity on bone marrow cells.

Further experiments were carried out to test whether the 'Epo-like' activity in SCM was: (1) Epo released by the spleen cells, (2) a factor present in SCM which enhanced the activity of undetectable amounts of Epo in the serum used to prepare the bone marrow cultures or (3) another factor(s) distinct from Epo, which was also able to stimulate late erythroid differentiation. SCM gave a similar dose-response curve whether 20% (v/v) or no serum was included in the bone marrow culture medium (Fig. 1), and the number of colonies formed was not related to the amount of serum added (data not shown). These results suggest that SCM is not enhancing the activity of any Epo which might be present in the culture medium.

Does SCM contain Epo or another factor(s)? To answer this question a comparison was made between the physical properties of the 'Epo-like' activity in SCM and those of Epo. The biological activities of both Epo and SCM were unaffected by dialysis against 0.16 M NaCl (Amicon PM 10 membrane; 10,000 MW cutoff), or by heating at 56 °C for 30 min, but both were rapidly inactivated by boiling at 100 °C. However, the SCM activity could be distinguished from Epo by column chromatography. Two peaks of activity were detected when SCM was passed over a Con A-Sepharose column (Fig. 2*a*)—the first small peak was eluted immediately after the void volume in a similar position to Epo. The second larger peak was eluted with α-methyl mannoside and coincided with the granulocyte/macrophage stimulating activity (GM-CSA). Two regions of activity were also observed after chromatography on Sephadex G-100 (Fig. 2*b*), the first of which ran with the serum proteins and had a MW similar to a bovine serum albumin (BSA) standard. The second was more pronounced, had an apparent MW of 20,000–40,000 and overlapped with the GM-CSA. In contrast to the results from separation on Con A-Sepharose, no activity was detected where Epo elutes from G-100 (Fig. 2*b*). Thus, both affinity and gel filtration chromatography demonstrate that SCM contains an erythroid colony-stimulating activity which is physically distinct from Epo but has a similar biological activity. Protein sequencing studies are necessary to determine whether this factor contains any structural homologies to Epo.

These results indicate that SCM contains a factor(s) capable of stimulating late erythroid differentiation in the absence of exogenous Epo. Factors other than Epo have previously been shown to influence early erythropoiesis[2,3] and in this respect it is interesting that Burgess *et al.*[13] have recently isolated a glycoprotein (apparent MW 37,000 by Sephadex G-150 chromatography; 24,000 following neuraminidase treatment) which induces 7 d erythroid colonies in fetal liver cell cultures in the apparent absence of Epo. The authors concluded that this molecule was similar to burst-promoting activity[2,3]: it possessed no Epo-like activity[13] and earlier studies showed that only selected batches of serum supported the development of haemoglobinized 7 d erythroid colonies[11] (suggesting that additional factors were required for the complete maturation of erythroid precursor cells). Thus, the present study is the first to demonstrate that Epo may not be the sole regulator of the late stages of erythroid development. Further experiments are required to determine the physiological significance of this factor *in vivo* and its relationship with Epo and other growth factors in the regulation of haematopoiesis.

I thank Dr N. Odartchenko for his support and encouragement, Dr H. Engers for help in preparing the SCM and Dr G. R. Johnson for PWM-SCM. This work was supported by an EMBO Research Fellowship, the Swiss NSF and the Schaffhauser Liga für Krebsbekampfuug.

Received 27 August; accepted 23 October 1980.

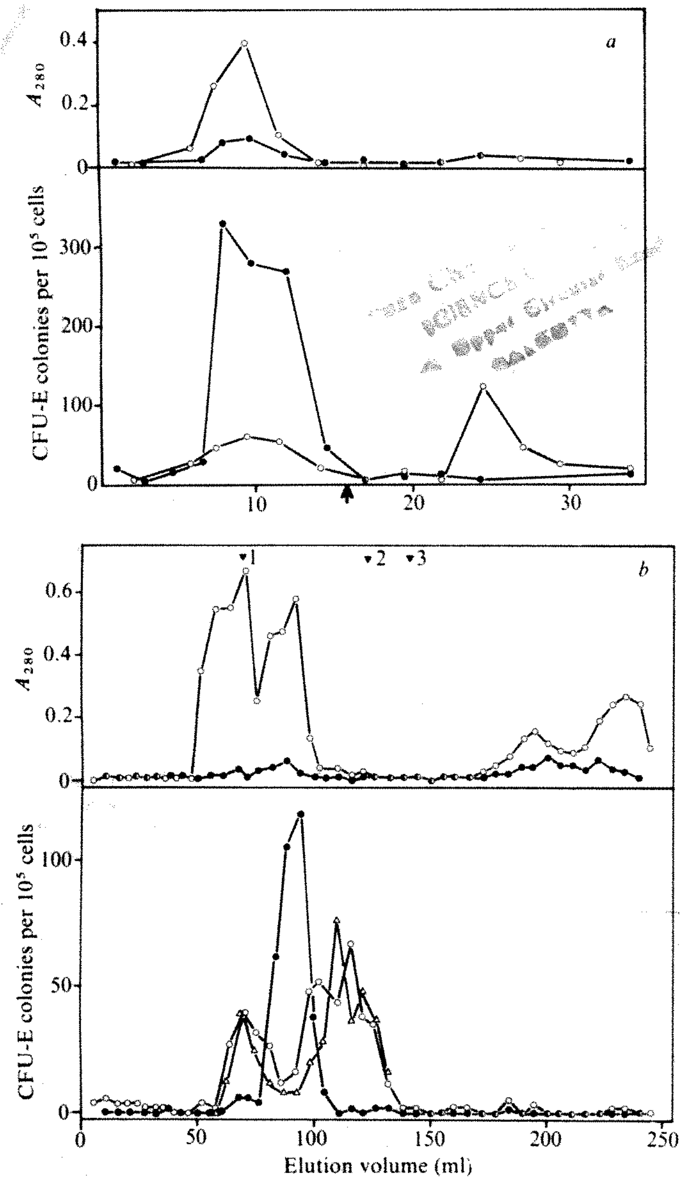


Fig. 2 Erythroid colony stimulating activity in fractions obtained by chromatography of Epo (●) and SCM (○, △) on (*a*) ConA-Sepharose and (*b*) Sephadex G-100 columns. The upper panel of (*a*) and (*b*) shows the A_{280} for each fraction: SCM contained 2% serum which accounts for the major peaks on these profiles. The lower panel shows the number of 2 d erythroid colonies formed in bone marrow cultures containing a 0.2 ml sample of each fraction. In all cases, the erythroid colonies were indistinguishable from those in control cultures containing Epo. *a*, ConA-Sepharose was equilibrated with running buffer (0.05 M sodium acetate, 0.5 M NaCl, 0.5 mM each of $MgCl_2$, $MnCl_2$ and $CaCl_2$, *p*H 6.5) and material bound to the column was eluted by addition of 0.1 M α-methylmannoside to the buffer (indicated by an arrow). Each fraction collected from the Con A column was dialysed against 0.16 M NaCl before assay. Similar results were obtained with other batches of SCM. *b*, Sephadex G-100 fine was equilibrated with 0.16 M NaCl, 4 mM HEPES, *p*H 7.4 containing 0.005% polyethylene glycol, which was also used to elute the samples. The increased A_{280} in the low-molecular weight region of the Sephadex G-100 profiles is due to phenol red which was present in the samples. Arrows 1, 2 and 3 indicate the positions of the MW standards, BSA (MW 69,000), α-chymotrypsin (MW 25,000) and ribonuclease (MW 13,700). The lower panel shows biological activity–elution profiles from two batches of SCM (○, △).

1. Krantz, S. B. & Jacobson, L. O. in *Erythropoietin and the Regulation of Erythropoiesis*, 13–24 (University of Chicago Press, 1970).
2. Iscove, N. N. & Guilbert, L. J. in *In Vitro Aspects of Erythropoiesis* (ed. Murphy M. J.) 3–7 (Springer, New York, 1978).
3. Wagemaker, G. in *In Vitro Aspects of Erythropoiesis* (ed. Murphy M. J.) 44–57 (Springer, New York, 1978).

4. Metcalf, D. *Recent Results Can. Res.* **61,** 161 (1977).
5. Iscove, N. N., Guilbert, L. J. & Weyman, C. *Expl Cell Res.* **126,** 121–126 (1980).
6. Parker, J. W. & Metcalf, D. *Immunology* **26,** 1039–1049 (1974); *J. Immun.* **112,** 502–510 (1974).
7. Metcalf, D., MacDonald, H. R., Odartchenko, N. & Sordat, B. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1744–1748 (1975).
8. Fagg, B. & Odartchenko, N. *Expl Hemat.* **7,** Suppl. 6, 123 (1979).
9. Ryser, J.-E., Cerottini, J-C. & Brunner, K. T. *J. Immun.* **120,** 370–377 (1978).
10. Johnson, G. R. & Metcalf, D. *Proc. natn. Acad. Sci. U.S.A.* **74,** 3879–3882 (1977).
11. Metcalf, D. & Johnson, G. R. *J. cell. Physiol.* **96,** 31–42 (1978).
12. Iscove, N. N., Sieber, F. & Winterhalter, K. H. *J. cell. Physiol.* **83,** 309–320 (1974).
13. Burgess, A. W., Metcalf, D., Russell, S. H. M. & Nicola, N. A. *Biochem. J.* **185,** 301–314 (1980).

Inhibition of histamine secretion from mast cells

M. Ennis, A. Truneh, J. R. White & F. L. Pearce

Department of Chemistry, University College London,
20 Gordon Street, London WC1H 0AJ, UK

Histamine secretion from mast cells may be inhibited by elevated intracellular levels of cyclic AMP and by several anti-allergic drugs[1–3]. These compounds are claimed to act directly on the calcium-gating mechanism activated by the anaphylactic reaction, preventing influx of Ca^{2+} from the external environment and so blocking exocytosis[4–7]. To examine this hypothesis further, we have compared here the histamine secretion induced by immunoglobulin E-directed ligands in the presence and absence of added calcium and by the ionophore A23187. Exocytosis evoked by these former agents was originally considered to be almost totally dependent on extracellular calcium[8,9] but recent studies have shown otherwise[10–14]. In the absence of added cation, the agents act by mobilizing membrane-bound or intracellular stores of calcium[10–12]. We show here that a variety of anti-allergic drugs are potent inhibitors in the conditions used, suggesting that alternative explanations for their action must be sought.

Rat peritoneal cells were recovered and histamine release determined as previously reported[10]. The effects of cromoglycate, theophylline and 8-bromo cyclic AMP on histamine release induced by antigen and concanavalin A (Con A) are shown in Table 1. The secretagogues produced similar releases of histamine in the presence and absence of added calcium[11,12] and, as in our previous work with the basic inducers, peptide 401, compound 48/80 and polylysine[11,15–17], these drugs were comparably effective in both situations. The inhibitory effects of quercetin and doxantrazole were also essentially independent of added calcium.

Two important differences between ionophore- and antigen-induced histamine release must be considered. First, high concentrations of ionophore can induce much larger releases of histamine than the anaphylactic reaction. Second, the signal produced by antigen is transient and decays rapidly with time, whereas that evoked by the ionophore is much more persistent[18].

The effects of cromoglycate and quercetin on the release induced by maximal and submaximal concentrations of ionophore are shown in Table 2. The compounds were effective inhibitors at the lower concentration of ionophore, which produced a release comparable with that evoked by antigen, but ineffective against the higher concentration. Theophylline, doxantrazole and 8-bromo cyclic AMP also inhibited histamine release induced by suboptimal amounts of ionophore[19]. These results are unlikely to reflect a simple interaction between the test drugs and the ionophore for several reasons. (1) The relative concentrations are inconsistent with such a model: cromoglycate (10 μM) inhibited significantly the release induced by the lower concentration of ionophore, whereas a 100-fold excess of drug was ineffective against a six-fold increase in the amount of

Table 1 Inhibition of histamine release induced by antigen and Con A in the presence and absence of extracellular calcium

	% Inhibition of release induced by:			
	Antigen		Con A	
	(+) Calcium	(−) Calcium	(+) Calcium	(−) Calcium
a, Cromoglycate (μM)				
200	—	—	55.1 ± 10.7	85.3 ± 6.8
100	61.7 ± 5.3	64.0 ± 10.0	51.5 ± 7.3	73.1 ± 6.4
50	62.9 ± 4.9	60.1 ± 10.1	46.8 ± 8.1	69.5 ± 7.3
25	57.5 ± 5.2	53.8 ± 10.0	42.6 ± 7.5	55.4 ± 7.3
10	43.2 ± 7.9	45.6 ± 10.2	31.5 ± 7.4	38.0 ± 6.6
5	22.8 ± 5.0	30.1 ± 13.3	25.0 ± 6.4	20.3 ± 3.2
1	6.3 ± 7.9	17.4 ± 6.9	8.9 ± 3.9	7.9 ± 1.2
b, Theophylline (mM)				
10	91.3 ± 6.7	95.4 ± 1.6	96.3 ± 1.2	100 ± 0.0
5	84.9 ± 6.2	77.2 ± 5.1	87.5 ± 3.5	85.4 ± 2.3
2.5	76.4 ± 8.1	57.7 ± 5.5	68.8 ± 8.7	57.7 ± 1.8
1	41.6 ± 8.0	30.8 ± 6.0	43.0 ± 7.7	31.4 ± 1.9
0.1	11.2 ± 4.8	−0.8 ± 2.6	2.6 ± 0.0	2.7 ± 1.1
c, 8-bromo cyclic AMP (mM)				
10	83.1 ± 4.1	96.0 ± 3.9	73.3 ± 3.2	80.4 ± 2.9
1	17.1 ± 8.3	23.0 ± 12.4	14.6 ± 1.8	−2.2 ± 9.8
0.1	4.6 ± 2.5	8.2 ± 8.2	1.9 ± 2.2	−7.8 ± 6.4

Rats were sensitized to the nematode *Nippostrongylus brasiliensis* and specific antigen was prepared as described previously[11]. Cells (1 ml) were preincubated, as previously reported[15,17], in HEPES-Tyrode buffer[10] containing calcium (1 mM) or EDTA (10^{-4} M) and then challenged with antigen (10 worm equivalents ml^{-1}) or Con A (20 μg ml^{-1}). Brief preincubation with EDTA enhances the secretory response in the absence of extracellular calcium, possibly by removing the cation from regulatory sites in the membrane and facilitating the release of more internal calcium stores[11,12]. Secretion was allowed to proceed for a further 10 min, terminated by addition of ice-cold buffer (2 ml) and assessed as before[10]. Values are means ± s.e.m. Control releases induced by antigen were: *a*: (+) calcium 30.2 ± 5.3, *n* = 6, (−) calcium 32.4 ± 2.7, *n* = 6; *b*: (+) calcium 35.9 ± 5.0, *n* = 4, (−) calcium 34.6 ± 4.6, *n* = 4; and *c*: (+) calcium 29.5 ± 0.9, *n* = 3, (−) calcium 31.2 ± 3.2, *n* = 3. Control releases induced by Con A were: *a*: (+) calcium 39.0 ± 3.9, *n* = 4, (−) calcium 31.5 ± 2.6, *n* = 4; *b*: (+) calcium 38.5 ± 3.8, *n* = 4, (−) calcium 44.8 ± 0.8, *n* = 4; and *c*: (+) calcium 57.1 ± 1.7, *n* = 4, (−) calcium 30.3 ± 3.0, *n* = 3.

Table 2 Inhibition of histamine release induced by different concentrations of ionophore A23187

	% Inhibition of release induced by ionophore at a concentration of:	
	1.5×10^{-7} M	10^{-6} M
a, Cromoglycate (μM)		
1,000	63.5 ± 3.3	8.4 ± 4.5
500	55.5 ± 5.9	5.6 ± 2.8
250	50.2 ± 7.6	1.1 ± 1.4
100	42.4 ± 7.8	0.0 ± 0.9
50	33.5 ± 7.0	0.4 ± 1.0
25	32.8 ± 8.2	1.3 ± 0.1
10	25.1 ± 7.1	−1.5 ± 1.4
b, Quercetin (μM)		
25	87.7 ± 2.8	18.5 ± 6.6
10	75.0 ± 12.1	6.4 ± 0.9
5	58.4 ± 13.0	3.7 ± 1.0
1	36.1 ± 10.7	2.2 ± 0.6

Cells were preincubated with quercetin (5 min, 37 °C) and stimulated with the ionophore. Cromoglycate was added with the releaser. Secretion was assessed after a further 10 min of incubation. Values are means ± s.e.m. Control releases induced by ionophore at 1.5×10^{-7} M were: *a*, 36.0 ± 5.7, *n* = 11, and *b*, 14.7 ± 6.3, *n* = 5; and at 10^{-6} M were: *a*, 77.9 ± 1.9, *n* = 3; and *b*, 73.7 ± 2.2, *n* = 5.

0028-0836/81/030186—02$01.00

Table 3 Effect of cromoglycate (200 μM) on cells stimulated with ionophore A23187 (10^{-6} M) for different times

Time (s)	Control histamine release (% of total)	% Inhibition of release by cromoglycate
10	1.2±0.4	—
20	32.3±4.1	94.3±1.5
30	41.7±3.3	91.0±2.4
40	44.5±4.2	79.4±3.5
60	48.8±5.4	58.3±4.2
90	49.2±6.2	40.9±4.4
120	52.1±7.2	29.6±10.7
300	69.3±2.0	18.1±2.7
600	75.3±2.7	9.0±5.9

Cells were stimulated with ionophore (with and without cromoglycate) for the period shown and the reaction terminated as in Table 1. Values are means ±s.e.m. for four experiments.

ionophore (Table 2). (2) The postulated interaction would have to apply to all the named compounds which represent diverse chemical types. (3) The drugs were active against maximal or supramaximal concentrations of ionophore if the cells were stimulated only for short periods of time (see below) or in the absence of extracellular calcium[19]. The concentrations required to effect inhibition in the latter conditions were generally appreciably less than in the presence of the cation and similar to those required to prevent antigen-induced secretion[19]. (4) This phenomenon is not confined to the ionophore and comparable effects (in which the extent of inhibition is dependent on the strength of the releasing stimulus) have been noted for various secretagogues[11,15–17].

To examine the time dependence of the inhibition, cells were stimulated with a maximal concentration of ionophore (with and without cromoglycate) and the reaction terminated after fixed periods. The control release of histamine increased progressively during the experiment but was potently inhibited by cromoglycate during the early stages of the reaction (Table 3). This effect is unrelated to the magnitude of the secretory response, as between 30 and 90 s of incubation there was only a slight increase in the amount of histamine released but a dramatic decline in the inhibitory effect of the drug. Thus, cromoglycate is a most potent inhibitor of exocytosis from cells stimulated with high concentrations of the ionophore for very short periods, conditions which mimic the transient signal produced by antigen. Identical results were obtained with doxantrazole and quercetin[19].

The tested anti-allergic compounds therefore effectively prevent the release of histamine in conditions where the calcium-gating mechanism is circumvented (either by stimulating the cells in the absence of extracellular calcium or with the ionophore in conditions more closely resembling the anaphylactic reaction) and alternative explanations for their action should be sought. The compounds might modulate the calcium signal by activating membrane pumps to extrude the cation from the cytosol or promote sequestration of the ion into internal stores, as does cyclic AMP in other systems[1]. Moreover, cyclic AMP and the anti-allergic compounds used here have been shown to prevent the accumulation of radiolabelled calcium by stimulated mast cells[4,6]. This effect was attributed to a decreased influx of the ion (and taken as evidence for an inhibition of the gating mechanism) but is equally well explained by an increased efflux of the cation. Accumulation studies alone will not distinguish between these possibilities[1]. Alternatively, the drugs might have effects not immediately related to calcium homeostasis and might promote disaggregation of microtubules[20], be involved in the phosphorylation of proteins in the plasma or perigranular membranes and so prevent the fusional changes involved in exocytosis[14,21], or have an as yet ill defined stabilizing effect on the mast cell membrane. Work to distinguish between these possibilities is now in progress.

This work was supported by grants from the MRC, the Wellcome Trust and the World University Service. We also thank Fisons Pharmaceuticals for financial assistance, and acknowledge gifts of cromoglycate (Fisons), doxantrazole (Wellcome) and the ionophore A23187 (Lilly).

Received 23 May, accepted 13 November 1980

1 Berridge, M J *Adv Cyclic Nucleotide Res* **6**, 1–97 (1975)
2 Garland, L G , Green, A F. & Hodson, H F in *Handbook of Experimental Pharmacology* Vol 50/II (eds Born, G V R , Farah, A , Herken, H & Welch, A D) 467–530 (Springer, Berlin, 1980)
3 Pepys, J & Edwards, A M (eds) *The Mast Cell, its Role in Health and Disease* (Pitman, Tunbridge Wells, 1979)
4 Foreman, J C , Hallett, M B & Mongar, J L *J Physiol , Lond* **271**, 193–214 (1977)
5 Foreman, J C , Mongar, J L , Gomperts, B D & Garland, L G *Biochem Pharmac* **24**, 538–540 (1975)
6 Foreman, J C , Hallett, M. B. & Mongar, J L *Br J Pharmac* **59**, 473–474P (1977)
7 Fewtrell, C M S & Gomperts, B D *Biochim biophys Acta* **469**, 52–60 (1977)
8 Foreman, J C & Mongar, J L *J Physiol , Lond* **224**, 753–769 (1972)
9. Foreman, J C , Mongar, J L & Gomperts, B D *Nature* **245**, 249–251 (1973)
10 Atkinson, G , Ennis, M & Pearce, F L *Br J Pharmac* **65**, 395–402 (1979)
11 Pearce, F. L *et al* in *The Mast Cell, its Role in Health and Disease* (eds Pepys, J & Edwards, A M) 69–75 (Pitman, Tunbridge Wells, 1979)
12 Ennis, M , Truneh, A , White, J R & Pearce, F L *Int Archs Allergy appl Immun* **62**, 467–471 (1980).
13 Baxter, J H & Adamik, R *Biochem Pharmac* **27**, 497–503 (1978)
14 Sullivan, T J & Parker, C W *Am J Path* **86**, 437–463 (1976)
15 Ennis, M , Atkinson, G & Pearce, F L *Ag Actions* **10**, 222–228 (1980)
16. Ennis, M , Pearce, F L & Weston, P M *Br J Pharmac* **70**, 329–334 (1980)
17 Ennis, M & Pearce, F L *Int Archs Allergy appl Immun* **60**, 461–464 (1979)
18 Foreman, J C , Garland, L G & Mongar, J L in *Calcium in Biological Systems* (ed Duncan, C J) 193–218 (Cambridge University Press, 1976)
19 Pearce, F L & Truneh, A *Ag Actions* (in the press)
20. Gillespie, E & Lichtenstein, L M *J clin Invest* **52**, 2941–2947 (1972)
21 Theoharides, T C , Sieghart, W , Greengard, P & Douglas, W W *Science* **207**, 80–82 (1980)

Persistent parainfluenza virus shedding during isolation at the South Pole

Harold G. Muchmore*†, Alan J. Parkinson*, Joe E. Humphries*, E. Nan Scott*, Douglas A. McIntosh*, L. Vernon Scott*, Marion K. Cooney‡ & John A. R. Miles§

* Departments of Medicine and Microbiology, College of Medicine, University of Oklahoma and Veterans Administration Medical Center, Oklahoma City, Oklahoma 73104
‡ Department of Pathobiology, University of Washington, Seattle, Washington 98195
§ Wellcome Virus Laboratory, Tamavua Hospital, Suva, Fiji

Persistent parainfluenza virus shedding in healthy young adults occurred throughout the 8½-month winter isolation period at Amundsen–Scott South Pole Station during 1978. Two episodes of respiratory illness were observed after 10 and 29 weeks of complete social isolation. Throat swabs collected both routinely, and during each outbreak of respiratory illness, were directly inoculated into cell cultures. Parainfluenza virus types 1 and 3 were recovered from both symptomatic and asymptomatic subjects throughout the winter. No other viruses were obtained by these efforts. The presence of parainfluenza virus in these subjects long after the accepted incubation period for viral upper respiratory illness, and when the introduction of new virus to this community was impossible, suggests its persistence in man.

Upper respiratory viruses can be demonstrated just before the onset of coryzal symptoms and for a few days during convalescence. The incubation period of parainfluenza virus illness is short in both natural and experimental human infections and ranges from 2 to 8 days. In long-range epidemiological studies in children, parainfluenza viruses were recovered from daily throat swabs for a period of 3–10 days, with a median of 8 days[1]. Similar brief virus shedding periods are observed following

---

† To whom reprint requests should be addressed

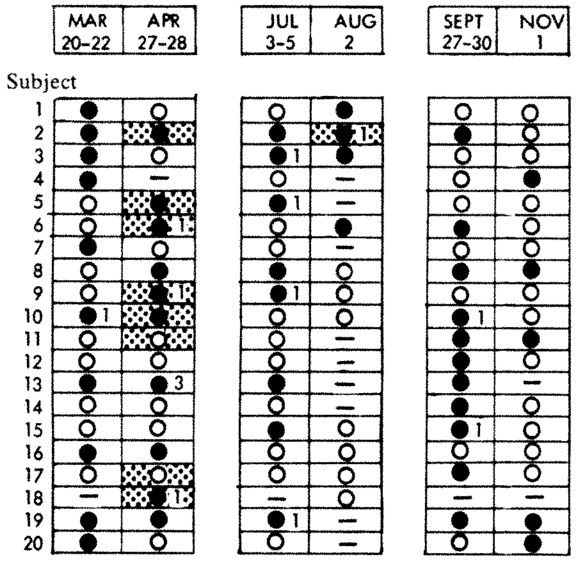


Fig. 1 Recovery of haemadsorbing viruses from throat swabs during the 1978 isolation period at the South Pole Station. Specimens were collected during the days shown for each month. The time interval between April and July and between August and September indicates that more than 30 days elapsed between specimens. ●, Virus recovered; ○, no virus recovered; —, no specimen obtained; ▨ indicates that the subject had an acute upper respiratory infection during this period; ▨ indicates that parainfluenza virus type 1 (or 3) was identified.

experimental infections in adults produced by nasal–oral inoculation of the parainfluenza virus[2]. In long-term studies of respiratory virus infections within open communities, persistent parainfluenza virus shedding has not been observed[3].

The American Antarctic Research Station at the geographic South Pole is totally isolated for 8½ months each year (mid-February to October) because the severe cold (−48 °C and below) prevents any aircraft operation. The summer period, extending from November to mid-February, is the time for station personnel changeover and re-supply. Each year since 1974, the South Pole physician has observed upper respiratory infections ('colds') among the Station personnel (~20 people) occurring after 2–6 months of total isolation, long after the usual short incubation period for these illnesses.

A limited virus laboratory was established at the South Pole Station, and operated during the austral winter of 1978 to attempt recovery and identification of the respiratory virus(es) responsible for these mid-isolation respiratory illnesses. Several continuous lines of mammalian tissue cells were maintained in culture by one of us (J.E.H.) throughout the year. These cell lines were human fetal tonsil, HeLa, Vero and Madin Darby canine kidney (MDCK). The last three lines were obtained as frozen suspensions from the American Type Culture Collection. The fetal tonsil line was isolated and tested by Schmidt *et al.*[4] and obtained as a confluent monolayer.

Cell monolayers were grown as needed on the inner surface of 2.0-ml glass screw-capped vials. The monolayers were nourished using Eagle's minimum essential medium (MEM) supplemented with 10% fetal calf serum (FCS) and antibiotics. Before inoculation, the growth medium was removed and replaced with an appropriate maintenance medium. For HeLa and fetal tonsil cells, Eagle's MEM with 2% FCS and antibiotics was used, but to support growth of influenza viruses and parainfluenza viruses in MDCK and Vero cells, a serum-free medium consisting of Eagle medium 199 supplemented with extra glucose, vitamins and 0.02% trypsin was used[5]. All cell-culture manipulations were carried out under a fibreglass microbiological hood equipped with a UV light source. Removal of either growth or maintenance medium from the cell cultures established in 2.0-ml vials was accomplished with a fresh, sterile, disposable, glass capillary tube connected to a closed-system vacuum source, and through suitable liquid traps, reducing the possibility of cross-contamination of cultures. Details of these miniature cell-culture methods were reported by Parkinson *et al.*[6].

The winter isolation period in 1978 began on 15 February and continued until 1 November. Routine throat swabs and serum specimens were obtained (after informed consent) in March, July and October 1978 from the 20 wintering personnel, all healthy young men. Throat swabs were swirled in a virus transport medium containing penicillin and streptomycin, and the resulting suspensions were inoculated into the monolayer vials and incubated (35 °C) on a roller drum apparatus. Uninoculated control monolayers of all cell lines were set up simultaneously and processed identically with each group of specimens collected. Acute upper respiratory symptoms were noted in a few persons in two episodes beginning on 22 April and 2 August 1978. At those times, additional monolayers were inoculated and processed along with additional controls for the presence of virus by standard virological methods[7]. After inoculation, the original throat-swab specimens were frozen and retained at −70 °C.

During the winter, cultures showing either cytopathic effect or haemadsorption were passaged into fresh monolayers, incubated and stored with the original cultures at −70 °C. After station opening on 1 November 1978, the stored inoculated cell cultures and the original specimens were reinoculated into new cell monolayers from fresh cell suspensions transported to South Pole Station from the US.

Many viruses were recovered by these efforts. The haemadsorbing agents were best detected by the Vero cells, although weak haemadsorption by some isolates could be detected in fetal tonsil cells. None of the control Vero and fetal tonsil cells showed any haemadsorption. No haemadsorbing or other viruses were detected in the HeLa and MDCK cell lines. From the 20 winter-isolation volunteers, haemadsorbing virus was obtained from 48 of 105 throat-swab specimens, a 40% recovery rate. The viruses were identified as parainfluenza virus types 1 and 3 by haemadsorption-inhibition techniques, using known virus-specific antiserum (National Center for Disease Control). Not all virus isolates were typed as the immediately available supply of typing serum was exhausted by the unexpectedly large number of viruses recovered.

Table 1 Serological responses of subjects wintering at South Pole Station during 1978 to parainfluenza virus types 1 and 3 as determined by haemagglutination inhibition

	Subject no.	Sep 1977	Mar 1978	Jul 1978	Oct 1978	Nov 1978
Parainfluenza type 1	5	160	160	20	80	40
	6	40	160	80	80	80
	11	20	10	10		80
	16	20	80	80	80	80
	20	20	20	40	160	80
Parainfluenza type 3	5	320	320	40	80	160
	7	160	80	160	40	40
	8		160	40	160	160
	11	160	40	10		160
	15	40	40	160	80	80
	19		80	320	160	160

Serological responses (underlined) are defined as a fourfold or greater change in haemagglutination-inhibition titre between successive serum specimens.

Virus recoveries throughout the isolation period are recorded in Fig. 1. Some subjects (nos 7, 12, 14, 17 and 18) yielded virus in only one sample, but most subjects shed virus on two or more occasions. Two subjects (nos 13 and 19) shed virus into each specimen collected throughout the winter isolation period. The figure also indicates those subjects experiencing symptoms of upper respiratory illness during the isolation period. Most of the virus shedders were asymptomatic throughout the winter.

After the frozen serum specimens reached the home laboratories, haemagglutination-inhibition tests were performed by standard methods[6,7] using commercial virus antigens (Microbiological Associates). Table 1 shows those subjects in whom four-fold changes of titre occurred. Subject numbers correspond to Fig. 1. Clearly, significant titre changes occurred in asymptomatic (nos 7, 8, 15, 16, 19 and 20) as well as symptomatic (nos 5, 6 and 11) individuals.

This remarkably high frequency of virus recovery indicates viral persistence and shedding throughout the isolation period in this small group of over-winter subjects. Such viral persistence in healthy subjects is contrary to present evidence that virus shedding in respiratory virus infections ends early in convalescence. However, in adult patients with chronic pulmonary disease, parainfluenza virus type 3 has been recovered on successive monthly occasions suggesting that this virus may persist in the abnormal respiratory tract[8]. The coryzal symptoms (colds) observed in the present group of subjects and in those of previous years occurred late in isolation, long after the presently accepted, short incubation period for parainfluenza virus infections. These late illnesses, occurring months into the total isolation period, also imply the continued presence of virus in an effective form which can be shed and cause disease. The South Pole environment is severely cold and dry. The mean ambient temperature is −47 °C, with a range of −18 °C. Only 1.5 cm of precipitation and 15 cm of new snow occur annually and the atmospheric water content (absolute humidity) averages 0.035 g m^{-3}. Whether this extremely cold, dry environment is significant in the observed parainfluenza virus persistence and shedding is unknown.

Allen *et al.*[9] reported a correlation between changes in outdoor temperature and onset of respiratory illness which occurred after 17 weeks of total isolation among members of the 1969 British Antarctic Survey party at Adelaide Island on the Antarctic Peninsula. In another study[10], subjects isolated at the British Antarctic Survey station, Stonington Island, for 5 months during 1971, showed an increased susceptibility to an experimental rhinovirus infection. An attenuated strain of virus was inoculated, but symptoms in subjects at Stonington Island were more severe than expected, with prolonged illness and virus shedding for a longer period (15 days) than observed in a control group inoculated with the same virus at the Common Cold Unit in Salisbury, England.

These earlier observations and our own reported here suggest that exposure of the upper respiratory tract to cold, dry air might impair the immune response of the respiratory mucosa. For example, this cooling and drying effect could lead to a reduced output of a secretory IgA and other anti-viral and anti-bacterial substances, resulting in either reactivation of an existing (persistent) viral infection or an increased susceptibility to reinfection.

This work was supported by NSF grant DPP 76-23789, the Oklahoma Medical Research Foundation and the Research Service of the Veterans Administration.

Received 14 July; accepted 12 November 1980.

1. Chanock, R. M., Bell, J. A. & Parrott, R. H. in *Perspectives in Virology* (ed. Pollard, M.) 128–138 (Burgess, Minneapolis, 1961).
2. Tremonti, L. P., Lin, J. L. & Jackson, G. G. *J. Immun.* **101,** 572–587 (1968).
3. Cooney, M. K., Fox, J. P. & Hall, C. E. *Am. J. Epidemiol.* **101,** 532–551 (1975).
4. Schmidt, O. W., Cooney, M. K. & Kenny, G. E. *J. clin. Microbiol.* **9,** 722–728 (1979).
5. Frank, A. L., Couch, R. B., Griffis, C. A. & Baxter, B. D. *J. clin. Microbiol.* **10,** 32–36 (1979).
6. Parkinson, A. J., Muchmore, H. G., Scott, L. V. & Miles, J. A. R. *J. clin. Microbiol.* **11,** 535–536 (1980).
7. Lennette, E. H. & Schmidt, N. J. in *Diagnostic Procedures for Viral and Rickettsial Infections,* 4th edn (American Public Health Association, 1969).
8. Gross, P. A., Green, R. H. & McCrea-Curnen, M. G. *Am. Rev. resp. Dis.* **108,** 894–898 (1973).
9. Allen, T. R., Bradburne, A. F., Stott, E. J., Goodwin, C. S. & Tyrrell, D. A. J. *J. Hyg., Camb.* **71,** 657–667 (1973).
10. Holmes, M. J., Reed, S. E., Stott, E. J. & Tyrrell, D. A. J. *J. Hyg., Camb.* **76,** 379–393 (1976).

Monoclonal anti-Fc receptor IgG blocks antibody enhancement of viral replication in macrophages

J. S. M. Peiris[*], S. Gordon[*], J. C. Unkeless[†] & J. S. Porterfield[*]

[*] Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK
[†] The Rockefeller University, 1230 York Avenue, New York, New York 10021

Flaviviruses, when complexed with antibody at subneutralizing concentrations, show enhanced replication in human and simian peripheral blood leukocytes (ref. 1, and J.S.M.P. and J.S.P., unpublished observations) and in P388 D1 and other macrophage cell lines[2]. A comparable phenomenon has been demonstrated with alphaviruses and Bunyaviruses in P388 D1 cells, (J.S.M.P. and J.S.P., unpublished observations) but cells lacking macrophage characteristics fail to show antibody-dependent enhancement (ADE) of viral replication. It has been suggested that the macrophage Fc receptor (FcR) provides an efficient route of entry of virus through the attachment of non-neutralized virus–antibody complexes and that for those viruses that escape destruction by the phagocyte, antibody results in a paradoxical increase in virus replication[3]. West Nile virus (WNV) replication in the P388 D1 macrophage cell line provides a reproducible model system for studying ADE of viral replication. Mouse macrophages have two FcRs—FcRI, which is trypsin-sensitive and binds IgG2a, and FcRII, which is trypsin-resistant and binds IgG2b and IgG1 complexes[4]. The FcR has been purified[5] using rat anti-mouse FcR monoclonal antibody which blocks FcRII[6]. We show here that anti-FcR IgG and its Fab fragment block ADE of virus replication by anti-WNV monoclonal antibodies.

The cultivation of P388 D1 cells[7] and preparation of WNV and rabbit anti-WNV serum have been previously described[3], and mouse monoclonal antibodies were prepared (J.S.M.P. and J.S.P., unpublished) by standard techniques[8]. Ascites fluids of clones producing anti-WNV antibody of immunoglobulin subclass IgG1 (F7/101) and IgG2a (F6/16A) were produced. Immunoglobulin subclass was determined by Ouchterlony double diffusion using two independent sources of subclass-specific antisera (Miles; Meloy). Both antiviral antibodies react with WNV antigen to high titre in radioimmunoassay and mediate enhanced plaque formation to a dilution of $10^{-5}-10^{-6}$ in P388 D1 cells. F7/101 neutralizes WNV poorly (50% plaque reduction at 1/10 dilution), whereas F6/16A neutralizes to a dilution of 1/1,280 (ref. 2). For the present experiments, both antibodies were used at subneutralizing concentrations (F7/101 at 1/1,000; F6/16A at 1/10,000). The production and characterization of the anti-mouse macrophage FcR monoclonal antibody and its Fab fragment have been described elsewhere[6].

As the anti-WNV antibody-induced enhancement of WNV yields in P388 D1 cells is correlated with the number of infectious centres (J.S.M.P. and J.S.P., unpublished observations), we used an infectious centre assay (as described in Fig. 1 legend) to quantify ADE of viral replication.

Figure 1 shows that addition of anti-WNV antibody F7/101 resulted in a 189-fold enhancement in the number of infectious centres formed. This effect was blocked by pretreatment of P388

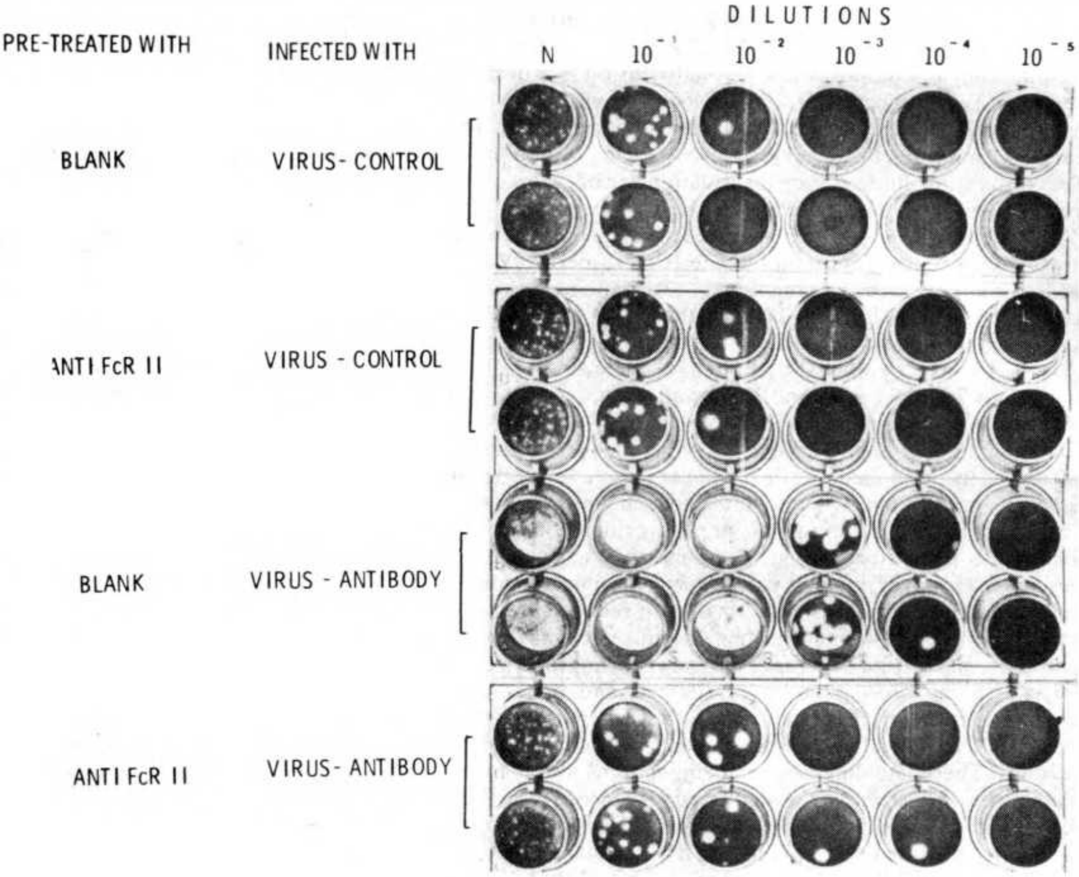


Fig. 1 Effect of anti-FcR antibody on ADE of WNV replication in P388 D1 cells. P388 D1 cells (2×10^6) were preincubated with anti-FcR antibody (240 μg ml^{-1}, 40 min, 4 °C) and infected with WNV (multiplicity of infection 0.04) with or without antiviral antibody F7/101 in the continued presence of anti-FcR (2 h, 4 °C). Cells were washed three times to remove unadsorbed virus and cell dilutions plated on a monolayer of pig kidney cells. An overlay of CM-cellulose was added 4 h later and plaques were stained after 4 days.

D1 cells with anti-FcR antibody, whereas virus replication in the absence of antiviral antibody was not affected. An irrelevant monoclonal antibody specific for mouse macrophages[9] had no effect in this system.

We next compared the inhibitory effect of IgG and the Fab fragment of anti-FcR antibody and examined the effect of antiviral antibody subclass (Fig. 2). Anti-FcR antibody (IgG) almost completely blocked ADE of virus replication induced by either anti-WNV IgG1 or IgG2a antibodies (Fig. 2*a*), in agreement with the ability of intact anti-FcR antibody to inhibit FcRI and FcRII in a macrophage rosetting assay[6]. In contrast, the Fab fragment of anti-FcR antibody showed less marked inhibition of ADE of virus replication, and no significant difference between the two subclasses (89% inhibition with F7/101 and 91% with F6/16A). The inhibition by anti-FcR Fab demonstrates that the inhibitory effect on ADE is due to the antibody combining site and not to the Fc domain. The reduced inhibitory activity of the monovalent Fab fragment may be due

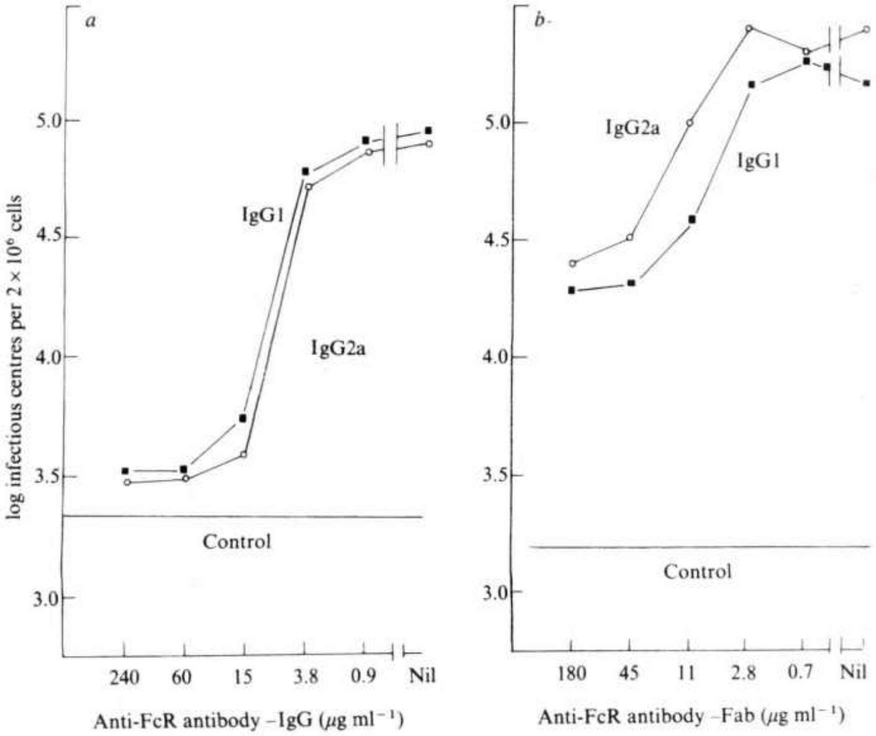


Fig. 2 Effect of the IgG (*a*) and Fab (*b*) fragments of anti-Fc receptor antibody on ADE of virus replication mediated by IgG1 and IgG2a anti-WNV antibodies. The procedures were as described in Fig. 1 legend.

Table 1 Comparison of the ability of anti-WNV IgG and F(ab')$_2$ to enhance replication of WNV

	WNV antibody titre			
Rabbit anti-WNV	Neutralization (50% plaque reduction)	Plaque enhancement	Radioimmune binding assay	mg protein per ml
IgG	1/40	1/2,000	1/3,125	0.86
F(ab')$_2$	1/10	<1/3	1/625	0.46

Anti-WNV IgG was purified by DE52 (Whatman) chromatography and digested with pepsin. F(ab')$_2$ fragments were isolated on Sephacryl S-200 (Pharmacia) and the absence of IgG confirmed by polyacrylamide gel electrophoresis. WNV neutralizing titre of antibody preparations was determined in pig kidney clone D cells[12]. Plaque enhancement was titrated in P388 D1 cell monolayers. The titre is taken to be the highest dilution giving at least threefold increase in plaque counts over controls[2]. For the radioimunoassay, WNV antigen or mock antigen was coated on polyvinyl microtitre plates (Dynatech) and incubated with serial dilutions of antibody. Antibody binding was detected using ^{125}I-labelled conjugated rabbit F(ab')$_2$ anti-mouse F(ab')$_2$. The antibody titre was the highest dilution of anti-WNV giving at least twofold greater binding to WNV antigen than mock antigen[13].

to the fact that binding to the FcR is reversible, whereas no measurable dissociation could be observed for the intact IgG[5]. The inability of Fab to discriminate between the IgG1 and IgG2 subclasses in ADE is surprising because previous studies showed that the same Fab fragment inhibited macrophage rosetting mediated by IgG1- and IgG2b-opsonized erythrocytes but was ineffective against IgG2a-mediated rosetting[6]. However, the binding of monomeric IgG2a to intact cells was inhibited by 15% by the anti-FcR Fab fragment and the binding of the purified FcR to IgG2a-coated beads could be completely inhibited[5].

To confirm that the Fc fragment of the antiviral antibody was responsible for ADE, bivalent F(ab')$_2$ pepsin fragments of anti-WNV rabbit antibody were prepared. The F(ab')$_2$ antibody lost its ability to enhance viral replication but retained substantial neutralizing activity and ability to bind to virus antigens on radioimmune binding assay (Table 1).

Several lines of evidence suggest that ADE of virus replication is involved in the pathogenesis of viral disease. Virus replication *in vivo* (influenza; dengue) increases in animals previously infected with antigenically related viruses or following transfer of low levels of antiviral antibody[10,11] and the dengue shock syndrome may be explained by this mechanism[3]. The present study shows that ADE of replication can be blocked by a monoclonal antibody specific for the Fc receptor. ADE demonstrated in human cultured peripheral blood monocytes can be abolished by removal of the Fc portion of the 'enhancing' antibody[1]. However, the presence of the Fc receptor, though necessary, is not sufficient to show ADE. Resident mouse peritoneal macrophages and the lymphoblastoid cell line WEH1 231 which possess the same FcR determinant as P388 D1 (ref. 6) do not show ADE (J.S.M.P. and S.G., unpublished). Where ADE is observed, it is not clear whether the FcR facilitates virus entry directly or whether, by concentrating virus on the cell surface, it increases infection through an independent receptor for the virus itself.

We thank J. Austyn for the monoclonal anti-macrophage antibody F4/80, A. F. Williams for providing us with rabbit anti-mouse F(ab')$_2$, and D. Millican and J. Wickson for technical assistance. This work was supported in part by the Leukaemia Society of America Inc., PHS grant AI-14603 and a grant from the MRC. J.C.U. is an American Cancer Society research scholar.

Received 11 September; accepted 13 November 1980.

1. Halstead, S. B. & O'Rourke, E. J. *J. exp. Med.* **146,** 201–217 (1977).
2. Peiris, J. S. M. & Porterfield, J. S. *Nature* **282,** 509–511 (1979).
3. Halstead, S. B. *Bull. Wld Hlth Org.* **58,** 1–21 (1980).
4. Diamond, B., Bloom, B. R. & Scharff, M. D. *J. Immun.* **121,** 1329–1333 (1978).
5. Mellman, I. S. & Unkeless, J. C. *J. exp. Med.* **152,** 1048–1069 (1980).
6. Unkeless, J. C. *J. exp. Med.* **150,** 580–596 (1979).
7. Snyderman, R., Pike, M. C., Fischer, D. G. & Koren, H. S. *J. Immun.* **119,** 2060–2066 (1977).
8. Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C. *Nature* **266,** 550–552 (1977).
9. Austyn, J. & Gordon, S. (in preparation).
10. Webster, R. G. & Askonas, B. A. *Eur. J. Immun.* **10,** 396–401 (1980).
11. Halstead, S. B. *J. infect. Dis.* **140,** 527–533 (1980).
12. Madrid, A. T. de & Porterfield, J. S. *Bull. Wld Hlth Org.* **40,** 113–121 (1969).
13. Rosenthal, J. D., Hayashi, K. & Notkins, A. L. *Appl. Microbiol.* **25,** 567–573 (1973).

Murine DNA repair gene located on chromosome 4

Pin-Fang Lin & Frank H. Ruddle

Department of Biology, Kline Biology Tower, Yale University,
New Haven, Connecticut 06511

Xeroderma pigmentosum (XP) is an autosomal recessive human disease in which affected individuals are prone to develop skin cancers after exposure to sunlight. Cells from XP patients are defective in DNA repair activity and are sensitive to UV light. Most XP individuals have a defect in the excision repair pathway for UV damage. Genetic studies have indicated that excision-defective cells fall into seven complementation groups (A–G). An eighth group of XP is known to be defective in post-replication repair[1–3]. DNA repair enzymes are ubiquitous and can function across species boundaries[4–6]. Primary mouse embryo fibroblasts have normal levels of DNA repair functions[7]. We report here that somatic cell hybrids between primary mouse cells and SV40-transformed XP group A cells can express wild-type levels of DNA repair function. These hybrid cells segregate murine chromosomes in culture. The proportion of cells in a given hybrid cell line which can perform unscheduled DNA synthesis (UDS) correlated well with the percentage of the population retaining murine chromosome 4. This study presents the first example of a direct quantitative comparison of specific gene activity and chromosomal content on a cellular basis.

Table 1 Statistical analysis of the relationship between retention of each murine chromosome and DNA repair activity in 21 hybrid lines

Mouse chromosome	Correlation coefficient	Slope	Intercept
1	0.422	0.446	34.1
2	0.445	0.451	34.3
3	0.486	0.587	22.4
4	0.940	0.992	2.01
5	0.599	0.625	28.9
6	0.554	0.775	32.6
7	0.485	0.789	34.4
8	0.611	0.538	28.6
9	0.564	0.620	28.9
10	0.438	0.617	30.6
11	0.527	0.768	28.9
12	0.391	0.390	26.5
13	0.612	0.805	28.1
14	0.558	0.721	33.0
15	0.740	0.774	16.8
16	0.575	0.623	25.9
17	0.658	0.614	25.0
18	0.401	0.413	22.0
19	0.364	0.435	27.7
X	0.451	0.676	34.4

The correlation coefficient *r* measures the degree of fit of the points on Fig. 2 (for chromosome 4), and similar plots for the other chromosomes, to the calculated least-square straight line. Slope indicates the number of murine chromosomes involved in conferring DNA repair ability on mouse–XP cell hybrids. Intercept represents residual DNA activity in the parental XP cells.

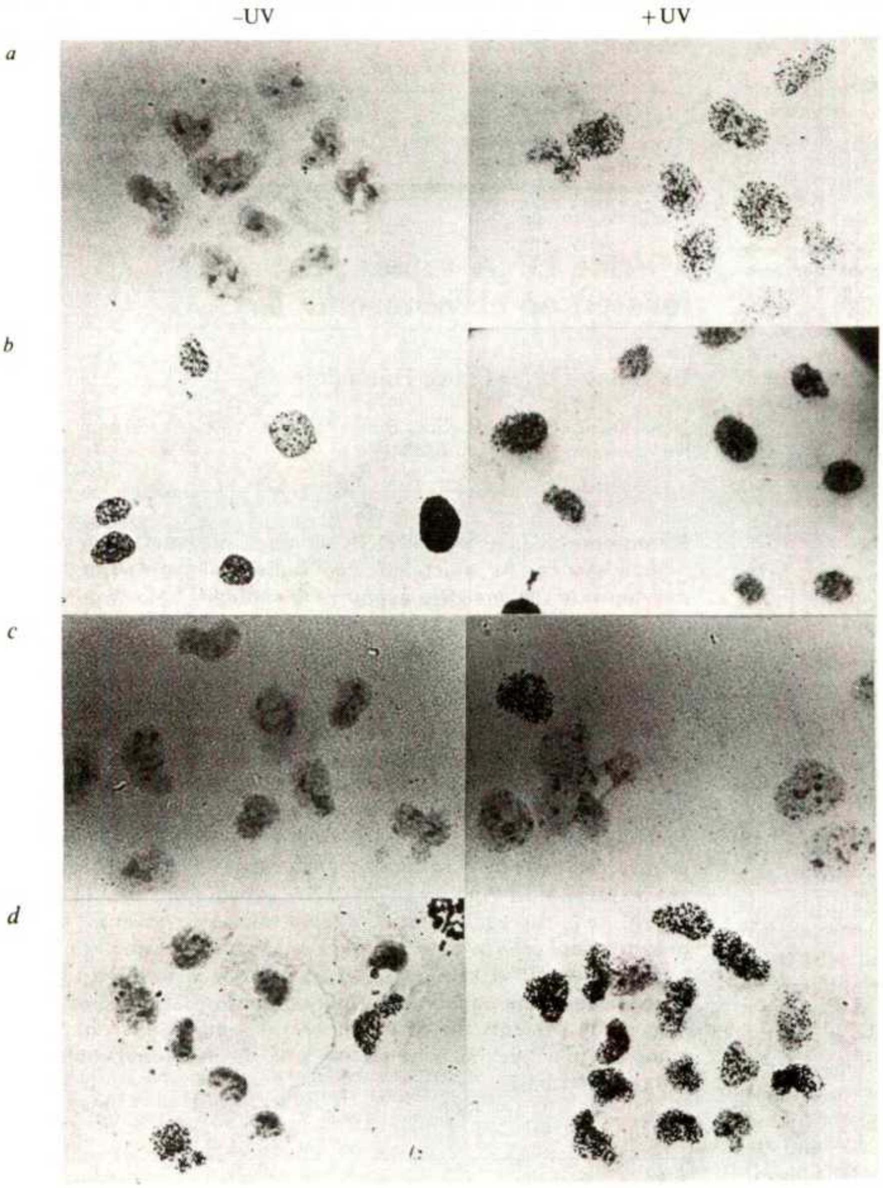


Fig. 1 Autoradiographs showing unscheduled DNA synthesis in parental and hybrid cell lines. Cells were UV-treated (30 J m^{-2}) and incubated in Dulbecco's modified Eagle medium + 10 μCi ml^{-1} ^{3}H-thymidine (specific activity 20 Ci $mmol^{-1}$) + 2 mM hydroxyurea for 2.5 h after irradiation. After washing and fixation, slides were exposed to Kodak NTB-2 emulsion at 4 °C for 7 days, developed then counterstained with Giemsa. Control cells were treated the same way, but were not UV-irradiated. *a*, SV40-transformed normal human cells, WI-18-VA-2 8-azaguanin resistant[11]. *b*, C57BL/6J mouse embryo fibroblast (MEF). *c*, SV40-transformed XP group A cells, 1223. *d*, Mouse–human hybrid cells $G_{31}B_2$ produced by fusion of MEF and 1223. The dark spots in unirradiated mouse nuclei (*b*) are murine chromocentres.

Sixteen independent cell hybrids were generated by fusing C57BL/6J mouse spleen and SV40-transformed XP group A cells and selecting for ouabain 10^{-6} M-resistant colonies[8] or by fusing microcells from mouse embryo fibroblasts and XP cells and selecting for UV (10 J m^{-2})-resistant colonies. The XP group A skin fibroblast (XP 12BE) was obtained from the American Type Culture Collection. Cell fusion was promoted by 50% polyethylene glycol (PEG) 1540 (ref. 9) or concanavalin A (Con A) and PEG (A. Lalazar, personal communication). The hybrids were tested for DNA repair by assay for unscheduled DNA synthesis (UDS) after UV-irradiation[10]. As shown in Fig. 1, cell hybrid $G_{31}B_2$ which contains a full complement of human and mouse chromosomes expresses a normal level of DNA repair activity. This result provides another example of the lack of species specificity for DNA repair. However, we cannot determine which species contributes the active DNA repair enzyme. The functional enzyme may be the mouse form, or activation of the human form may occur in the presence of mouse regulating elements.

It has been shown that after several successive transfers of primary mouse embryo fibroblasts *in vitro*, the ability of these cells to perform excision repair decreases[7]. Note that in the stable hybrid $G_{31}B_2$ there is no apparent change in DNA repair capacity after 20 months in culture.

Most mouse–XP hybrids obtained segregated mouse chromosomes and contain an average of 60 human chromosomes per cell. Chromosome analyses[12] and UDS assays were performed to determine whether there is an association between the segregation of a specific mouse chromosome and the expression of DNA repair function in mouse–XP cell hybrids. To confirm the karyotypic results, cell extracts from each hybrid were also assayed for the presence of 18 murine isozymes which have been mapped (by starch gel electrophoresis) in the mouse genome to 13 chromosomes[13–15].

The DNA repair activity of cell hybrids was tested by UDS and the percentage of cells which could perform DNA repair calculated as follows:

$$\frac{\left(\begin{array}{c}\text{\% cells incorporating}\\ ^3\text{H-thymidine after UV}\end{array}\right)-\left(\begin{array}{c}\text{\% cells incorporating}\\ ^3\text{H-thymidine without UV}\end{array}\right)}{1-(\text{\% cells incorporating } ^3\text{H-thymidine without UV})}$$

This represents:

$$\frac{\left(\begin{array}{c}\text{\% irradiated cells in S}\\ \text{phase and in DNA repair}\end{array}\right)-\left(\begin{array}{c}\text{\% control cells}\\ \text{in S phase}\end{array}\right)}{1-(\text{\% control cells in S phase})}$$

In the experimental conditions 50% of unirradiated SV40-transformed normal human fibroblasts (WI-18-VA-2) and SV40-XP A cells were found to be engaged in DNA replication. After UV irradiation (30 J m^{-2}) ~ 98% of the WI-18-VA-2 cells showed DNA repair activity, whereas the value for SV40-XP A was only 2%.

In this study we analysed 21 hybrid lines (16 primary clones and 5 subclones), which cover the full complement of murine chromosomes except Y. To study the effect of each donor murine chromosome on the expression of DNA repair function in XP cell hybrids, the percentage of cells expressing DNA repair activity was plotted against the percentage of cells retaining a particular murine chromosome for the 21 hybrid lines. Figure 2 shows an example with mouse chromosome 4. For each of the 20 mouse chromosomes a least-square linear regression analysis of these two parameters yielded the results shown in Table 1.

The correlation coefficient r measures the degree of fit of the given points to the least-square straight line and the slope indicates the number of murine chromosomes which are involved in complementing the DNA repair defects in XP cells. The intercept represents the residual DNA repair activity in the parental XP cells which is about 2% (determined experimentally). The analysis shows that the hybrid cells which contain murine chromosome 4 have the highest correlation coefficient value (0.94) and the correct intercept value (2.01). Furthermore, in a Student's t test, for $r = 0.94$, $P < 0.001$. These data indicate that mouse chromosome 4 is responsible for the expression of DNA repair function in XP cell hybrids. The slope of 0.992 further suggests that murine chromosome 4 alone can complement the XP group A defect(s).

Another statistical analysis of expression of DNA repair function and karyotypic data for the hybrid line studies was carried out using a computer program[16] and the results are summarized in Table 2. The positive correlation between the presence of an individual chromosome in a clone and that clone's ability to express DNA repair function was computed with three different measures of dependence. In all three cases, the expression of DNA repair activity was correlated with the presence of chromosome 4.

Table 3 lists the frequency of retention for mouse chromosome 4 and the percentage of cells which regain UDS for five subclones which were derived from four independent hybrids. These data further indicate that a change in the retention frequency of murine chromosome 4 is accompanied by a parallel change in the DNA repair ability in subclone lines.

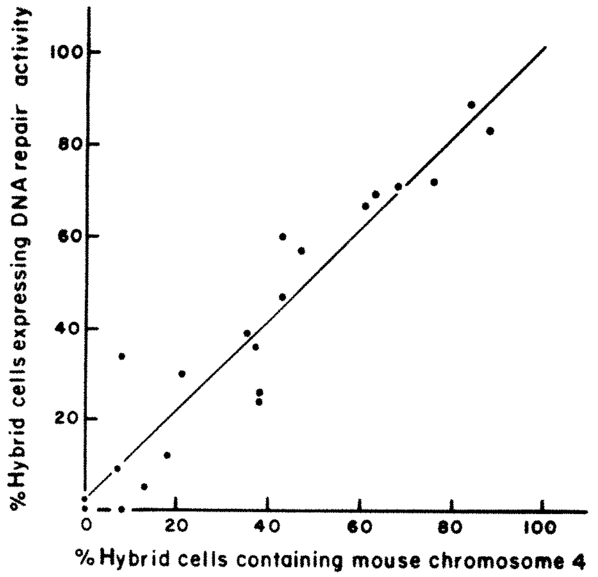


Fig. 2 The relationship between the presence of murine chromosome 4 and DNA repair activity in 21 murine–XP hybrid lines. Cells were treated as in Fig. 1 legend. Over 500 cells were counted for the incorporation of ^{3}H-thymidine for each hybrid. Cells were regarded as unlabelled if there were ten grains or less per nucleus.

Table 2 Assignment of murine gene responsible for restoring DNA repair activity in mouse–XP group A cell hybrids to chromosome 4

Chromosome	UDS/chromosome (no. of clones)				ϕ	χ^2	Overall OR
	+ +	+ −	− +	− −			
1	4	11	2	4	−0.07	0.09	2.69
2	4	11	1	5	0.11	0.24	2.20
3	11	4	3	3	0.22	1.05	7.38
4	14	1	0	6	0.89	16.80	10.56
5	6	9	1	5	0.22	1.05	3.45
6	3	12	1	5	0.04	0.03	1.90
7	3	12	0	6	0.26	1.40	2.19
8	7	8	0	6	0.45	4.20	5.03
9	6	9	2	4	0.06	0.08	2.92
10	6	9	1	5	0.22	1.05	3.50
11	6	9	0	6	0.40	3.36	3.87
12	11	4	4	2	0.07	0.09	7.60
13	5	10	0	6	0.35	2.62	2.97
14	4	11	0	6	0.31	1.98	2.44
15	11	4	2	4	0.37	2.91	7.26
16	8	7	1	5	0.33	2.35	6.38
17	7	8	2	4	0.12	0.31	4.58
18	11	4	3	3	0.22	1.05	6.91
19	10	5	5	1	−0.17	0.58	6.54
X	2	13	2	4	−0.23	1.11	1.37

Hybrid cells were considered positive for DNA repair function when over 20% of cells expressed UDS. For chromosomes (+) indicates presence of at least one copy in over 20% of the cells analysed. Statistical analysis was carried out using a computer program, ASSIGN, developed for the purpose of gene assignment with material generated by somatic cell hybridization[16]. To determine the gene assignment, three different statistical procedures were used: χ^2, ϕ and OR. χ^2 was used as a method of dependence, rather than a test of significance, and was calculated in the ordinary way. The ϕ coefficient, which is insensitive to sample size is the ratio of the observed χ^2 to the perfect χ^2. It is calculated by dividing the square root of the regular χ^2 by the square root of the sample size. When there exists a perfect association between a marker and a chromosome, the value of ϕ is +1. Zero indicates no association and −1 a negative association. The OR statistic was designed for determining gene assignments. It is an inverse weighted scoring system, independent of χ^2, which is calculated by assigning each clone a score of 100 and dividing that value evenly among all chromosomes in concordance with the tested marker. The clones are scored, the scores added for each chromosome, and then divided by the number of clones. The chromosome with the highest number is considered a candidate for assignment of the gene under study. Generally an OR score of 7 or above is considered evidence for a positive assignment.

Table 3 The concordant segregation of chromosome 4 and the expression of DNA repair function in subclones

Hybrid line	Frequency of chromosome 4	% Cells with UDS
1. $S \times O_2$	61	67
$S \times O_2 - 1$	37	36
$S \times O_2 - 50$	76	72
2. $S \times A_2$	88	83
$S \times A_2 - 1$	21	30
3. $S \times R5$	35	39
$S \times R5 - 1$	43	47
4. $G_{31}B_2$	84	89
$G_{31}B_2 - 1$	43	60

Therefore we conclude that chromosome 4 contains a gene which corrects the defective DNA repair function in XP A cells. As it is not known if the XP A defect involves a structural gene change or an altered regulatory element, the gene product of chromosome 4 responsible for restoring DNA repair function remains undefined. Extending similar studies to other XP groups should help in elucidating the arrangement of murine DNA repair genes and mechanisms of mammalian DNA repair. The repair pathway for DNA damage in cells is complex, involving many different enzymatic activities; it is important to determine where the genes for these enzymes reside in the genome and whether they are closely linked.

We thank Dr James McDougall for SV40-transformed XP cells; Mrs F. Lawyer and Dr D. Pravtcheva for help with karyotypic analysis; Ms E. Nichols, Mr D. Bonito, Mr V. Salerno and Ms S. Pafka for technical assistance; Ms U. Shankar for performing computer analysis; Mrs M. Reger and Mrs M. Siniscalchi for manuscript preparation and Drs D. Slate and D. Juricek for helpful discussions and criticisms of the manuscript. This work was supported by USPHS grants 5R01/GM09966 and AGO1940 to F.H.R.

Received 12 September; accepted 6 November 1980.

1. Cleaver, J. E. & Bootsma, D. *A. Rev. Genet.* **9,** 19–38 (1975).
2. Artlett, C. F. & Lehman, A. R. *A. Rev. Genet.* **12,** 95–115 (1978).
3. Setlow, R. B. *Nature* **271,** 713–717 (1978).
4. Tanaka, K., Sekiguchi, M. & Okada, Y. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4071–4075 (1975).
5. Tanaka, K., Hayakawa, H., Sekiguchi, M. & Okada, Y. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2958–2962 (1977).
6. Darzynkiewicz, Z. & Chelmicka-Szore, E. *Expl Cell Res.* **74,** 131–139 (1972).
7. Peleg, L., Raz, E. & Ben-Ishai, R. *Expl Cell Res.* **104,** 301–307 (1976).
8. Kucherlapati, R. S., Baker, R. M. & Ruddle, F. H. *Cytogenet. & Cell Genet.* **14,** 362–363 (1975).
9. Davidson, R. L. & Gerald, P. S. *Somat. Cell Genet.* **2,** 165–176 (1976).
10. Goldstein, S. *Proc. Soc. Expl Biol. Med.* **137,** 730–734 (1971).
11. Weiss, M. C., Ephrussi, B. & Scaletta, L. J. *Proc. natn. Acad. Sci. U.S.A.* **59,** 1132–1135 (1968).
12. Kozak, C. A., Lawrence, J. B. & Ruddle, F. H. *Expl Cell Res.* **105,** 109–117 (1977).
13. Kozak, C. A., Fournier, R. E. K., Leinwand, L. A. & Ruddle, F. H. *Biochem. Genet.* **17,** 23–34 (1979).
14. Omenn, G. A. & Cohen, P. T. W. *In Vitro* **7,** 132–139 (1971).
15. Lin, P. F., Slate, D., Lawyer, F. & Ruddle, F. H. *Science* **209,** 285–287 (1980).
16. Cowmeadow, M. & Ruddle, F. H. *Cytogenet. & Cell Genet.* **22,** 694–697 (1978).

Inducible repair of near-UV radiation lethal damage in *E. coli*

J. Peters & J. Jagger

Biology Programs, The University of Texas at Dallas, PO Box 688, Richardson, Texas 75080

We present here the initial characterization of a new repair system which is induced in actively growing cultures of *Escherichia coli* and which repairs a major fraction of lethal damage produced by near-UV (310–400 nm) light. This system is different from the error-prone 'SOS' repair system for DNA, known to be induced in *E. coli* by treatments which damage DNA and/or inhibit DNA replication, such as irradiation by far-UV (254 nm) light or X rays, thymine starvation or treatment with naladixic acid or mitomycin C[1,2]. The SOS response requires induction of an 'X' protein[3], the product of the *recA* gene[4–6], whereas the inducible repair system described here utilizes proteins distinct from the X protein.

The induction of a near-UV repair system different from SOS repair was initially inferred from the unusual survival pattern of *E. coli* K12 AB 2463 observed after irradiation with 366-nm monochromatic light. This strain carries the *recA*-13 allele, which renders it incapable of inducing the X protein or showing the physiological responses associated with SOS repair[7]. 'V'-shaped survival curves were obtained when exponential-phase cells were irradiated in conditions permitting protein synthesis during irradiation (Fig. 1). After the initial reduction of viability with fluences up to 70–100 kJ m^{-2}, further irradiation resulted in recovery of colony-forming ability; that is, cells, which would have been non-viable when plated after lower fluences, were viable when allowed to remain in suspension for additional irradiation before plating. Recovery is maximal after 200–250 kJ m^{-2}, the cells remaining resistant up to fluences of 500 kJ m^{-2}. Although the extent of killing and recovery can be modified by post-irradiation conditions, the V-shaped survival pattern remains, suggesting that the critical repair event(s) occurs rapidly, during irradiation and before plating. Addition of the protein synthesis inhibitor, chloramphenicol (CAM), at the start of irradiation, results in nearly exponential loss of viability, with only a slight low-fluence shoulder (Fig. 2). However, when CAM is added after the cells have received 200 kJ m^{-2}, there is a pronounced shoulder of about 40 kJ m^{-2} (Fig. 2).

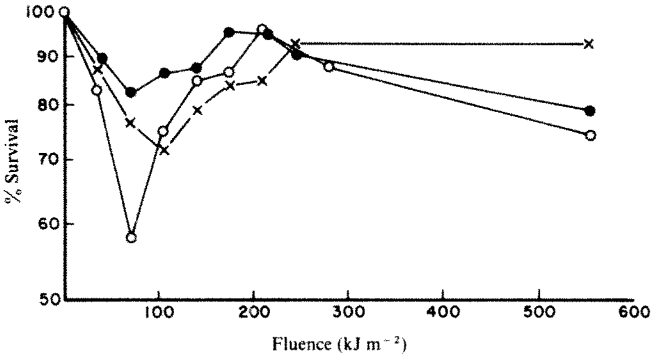


Fig. 1 Percentage survival of colony-forming ability of *E. coli* K12 AB 2463 (*recA*) as a function of 366-nm fluence in different post-irradiation conditions: plated on Luria agar (10 g l^{-1} tryptone; 5 g l^{-1} yeast extract; 5 g l^{-1} NaCl; 15 g l^{-1} agar) and incubated at 25 °C (×), plated on nutrient agar and incubated at 37 °C (○) or at 25 °C (●). Exponential-phase cultures were grown and irradiated (at 5–7.5 × 10^{7} cells ml^{-1}) at 37 °C in complete M-9 mineral salts medium[12] containing 0.4% glucose and supplemented with the auxotrophic requirements: arginine, histidine, threonine and leucine, each at 20 μg ml^{-1}, and thiamine at 5 μg ml^{-1}. The source of monochromatic 366-nm light was a Philips SP-500 high-pressure mercury arc lamp operated in conjunction with a Bausch and Lomb grating monochromator (model 33-86-45-49). A filter (Corning 0-52) was placed at the exit slit to screen out 340 nm and shorter-wavelength UV light. The fluence rate, measured with an Eppley thermopile, was 130 W m^{-2}.

One interpretation of these results is that recovery of viability occurring in conditions that allow protein synthesis during irradiation is due to the induction of a repair system not present in unirradiated cells, or in cells receiving less than the inducing fluence of about 70 kJ m^{-2}. CAM added at the start of irradiation prevents induction of the repair system, whereas the shoulder observed when CAM is added after 200 kJ m^{-2} reflects the time required for decay of induced repair protein(s) to

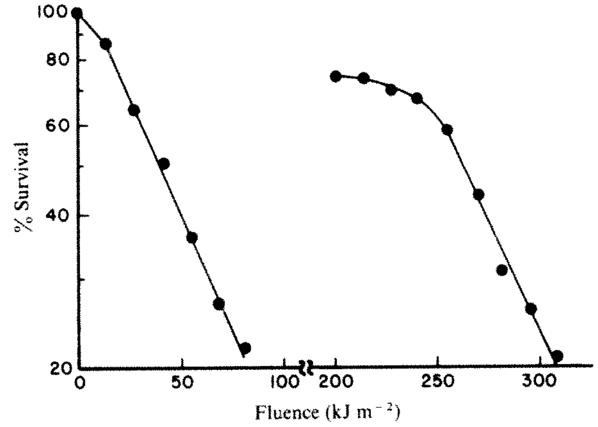


Fig. 2 The effect of inhibition of protein synthesis on colony-forming ability after near-UV (366 nm) irradiation. Cultures of *E. coli* K-12 AB 2463 were grown and irradiated as in Fig. 1 and plated on Luria agar at 25 °C. Chloramphenicol was added to the complete medium to a final concentration of 60 μg ml^{-1} at the start of irradiation (left-hand curve) or after a fluence of 200 kJ m^{-2} (right-hand curve). The fluence rate was 226 W m^{-2}. Complete medium irradiated with or without chloramphenicol had no phototoxic effect when incubated with unirradiated cells.

0028–0836/81/030194—02$01.00

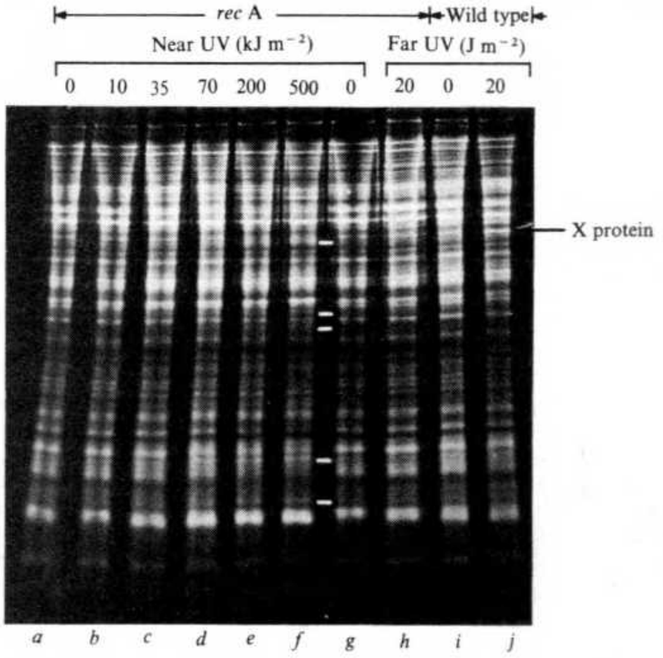


Fig. 3 Autoradiograph of SDS-polyacrylamide gel electrophoresis of total cellular protein of *E. coli* labelled with ^{35}S-methionine. The figure shows new protein bands appearing during near-UV (366 nm) irradiation of *rec*A cells and comparison of the near-UV-inducible proteins with the X protein induced after far-UV (254 nm) irradiation of wild-type cells. Proteins made in unirradiated *E. coli* K12 AB2463 *rec*A cells are shown in lanes *a* and g, and after increasing fluences of near-UV light in lanes *b–f*. The corresponding fluences are shown at the top of each lane. Proteins made in unirradiated wild-type *E. coli* K12 AB1157 cells are shown in lane *i* and in far-UV irradiated wild-type cells in lane *j* and *rec*A in lane *h*. Bands of interest are highlighted by white lines in the gap between lanes *f* and g. The X protein in lane *j* is indicated by the arrow. Growth and near-UV irradiation conditions are as in Fig. 1. Cultures (0.5 ml) were pulse-labelled for 90 s in medium supplemented with 0.5 μg ml^{-1} cold methionine and 50 μCi ml^{-1} of ^{35}S-methionine. Total cellular protein was solubilized and analysed by SDS-polyacrylamide gel electrophoresis using 14.5% gels according to the procedure of Studier[13]. Autoradiograms of dehydrated gels were made using Kodak X-Omat-R X-ray film. Pulse labelling of near-UV-irradiated cultures was done during continued irradiation; the far UV-irradiated culture was incubated for 20 min at 37 °C after irradiation before pulse labelling, to allow induction of the X protein.

physiologically insignificant levels. This period was estimated (from the ratio of the shoulder fluence to the fluence rate) to be 3.5 ± 0.5 min, and it remains constant over a threefold change in fluence rate. In contrast, the biological half life of SOS repair, as measured by the decay of error-prone repair activity, is 20–30 min (refs 8, 9).

This induced repair system must repair at least 90% of the damage produced at 366 nm, as may be seen by comparing the survival when the repair system is fully induced (>200 kJ m^{-2} fluence in Fig. 1) with that when its induction is prevented (left-hand curve in Fig. 2).

The interpretation of these survival data in terms of an inducible repair system is supported by our detection of new proteins synthesized during irradiation, whose appearance coincided with the recovery of colony-forming ability. Cultures of the *rec*A strain were pulse-labelled with ^{35}S-methionine during irradiation with 366-nm light, and total cellular protein was analysed by SDS-polyacrylamide gel electrophoresis followed by autoradiography. Lanes *a* and *g* of Fig. 3 show the pattern of pulse-labelled proteins from unirradiated *rec*A cells, and lanes *b–f* show the proteins from cells that were pulse-labelled after increasing fluences of 366-nm light, from 10 to 500 kJ m^{-2}. Three new protein bands (which we call P_1–P_3 in order of decreasing molecular weight, starting at the top of Fig. 3) appear at about 70 kJ m^{-2} (lane *d*), and the bands corresponding to proteins P_4 and P_5 (indicated by the bottom two white lines in Fig. 3) show large dose-dependent increases in comparison with the unirradiated control. The relative intensity of the bands increases up to about 200 kJ m^{-2} (lane *e*), at which time the recovery of viability is at a maximum (Fig. 1), with the pattern remaining essentially unchanged up to 500 kJ m^{-2} (lane *f*).

Pulse-labelled proteins from unirradiated wild-type parental strain K12 AB 1157 are shown in lane *i* of Fig. 3. Lane *j* shows the proteins from wild-type cells that were pulse-labelled 20 min after irradiation with 20 J m^{-2} of 254-nm far-UV light (the arrow indicating the X protein), whereas the same treatment failed to induce the X protein in the *rec*A strain (lane *h*). This is generally the case with strains carrying mutant *rec*A alleles in a variety of inducing conditions[7,10]. Proteins P_1–P_5, which are inducible in *rec*A cells by near-UV light, can be distinguished from the X protein by electrophoretic mobility, consistent with the idea of a *rec*A gene-independent, inducible repair system.

To distinguish near-UV-inducible repair from SOS repair, the above experiments were carried out in a *rec*A strain, which lacks SOS repair. Although the wild-type parental strain did not show V-shaped survival curves, the following results indicate that the near-UV-inducible repair pathway also makes a major contribution to survival of the wild-type strain after near-UV irradiation: the extensive shoulder characteristic of near-UV survival curves is greatly reduced in the presence of CAM; the fluence at 366 nm required to reduce survival of strain AB 1157 to 10% of the unirradiated controls is 3,500 kJ m^{-2} without CAM and 240 kJ m^{-2} when CAM is added to exponential cultures at the start of irradiation. Proteins P_1–P_5 are also induced in the wild-type strain by near-UV irradiation, as well as in the *rec*A strain.

The maximal solar UV-A (320–400 nm) irradiance is about 5 mW cm^{-2} at sea level[11]. In these conditions, a fluence of 200 kJ m^{-2} is delivered in about 1 h of exposure. Thus, the threshold fluence for the induction of the repair system described here is within the range of physiological exposure to UV-A, and we expect that this type of repair may be an adaptation to environments exposed to near-UV solar radiation.

This research was supported by USPHS grant AM 14893 and the University of Texas at Dallas Research Fund.

Received 10 July; accepted 7 November 1980.

1. Radman, M. in *Molecular Mechanisms for Repair of DNA* (eds Hanawalt, P. C. & Setlow, R. B.) 355–367 (Plenum, New York, 1975).
2. Witkin, E. M. *Bact. Rev.* **40,** 869–907 (1976).
3. Gudas, L. J. & Pardee, A. B. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2330–2334 (1975).
4. Gudas, L. J. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5280–5284 (1977).
5. Emmerson, P. T. & West, S. C. *Molec. gen. Genet.* **155,** 77–85 (1977).
6. McEntee, K. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5275–5279 (1977).
7. Gudas, L. J. *J. molec. Biol.* **104,** 567–587 (1976).
8. Witkin, E. M. *Molec. gen. Genet.* **142,** 87–103 (1975).
9. Defais, M., Caillet-Fuquet, C., Fox, M. S. & Radman, M. *Molec. gen. Genet.* **148,** 125–130 (1976).
10. Inouye, M. *J. Bact.* **106,** 539–542 (1971).
11. Parrish, J. A., Anderson, R. R., Urbach F. & Pitts, D. *UV-A: Biological Effects of Ultraviolet Radiation with Emphasis on Human Responses to Long-wave Ultraviolet*, 9–15 (Plenum, New York, 1978).
12. Davis, B. D. & Mingioli, E. S. *J. Bact.* **60,** 17–28 (1950).
13. Studier, F. W. *J. molec. Biol.* **79,** 237–248 (1973).

Evidence for the inducibility of the *uvrB* operon

Michael Fogliano & Paul F. Schendel

Genetics and Cell Biology Section, Biological Sciences Group, Institute of Materials Science, The University of Connecticut, Storrs, Connecticut 06268

A new tool for probing the regulation of non-essential gene products has recently been developed[1]—this is a derivative of bacteriophage Mu into which the *Escherichia coli* genes for lactose metabolism have been fused (referred to here as Mu(lac, amp)). The DNA of this defective virus can insert randomly into the chromosome of a cell and mutants can be retrieved in which the β-galactosidase gene carried on the phage is under the control of the regulatory elements of the operon in which it now resides. We used this phage to generate DNA repair-deficient mutants, one of which was highly UV-sensitive due to insertion of the Mu(lac, amp) into the *uvrB* region of the chromosome. When exposed to UV light, this strain produces β-galactosidase at a much greater rate than unexposed cells. This suggests that the *uvrB* operon is inducible. These results were initially rather surprising. Although various components of the enzyme systems involved in Weigle-reactivation[2,3], post-replication repair[4–6] and repair of alkylation damage[7,8] had been established as inducible, the products of the *uvrA*, *uvrB* and *uvrC* genes which collectively effect nucleotide excision repair[9–11], seemed to be firmly established as a constitutively synthesized repair system. This conclusion was drawn because the activity of these enzymes can be easily detected in uninduced cells[9–11]; and mutations in the genes *recA* and *lexA*, which affect the regulation of most of the inducible repair systems[12], have relatively little effect on *uvrA*, *uvrB*-mediated excision repair[13]. Thus the *uvrA* and *uvrB* gene products seemed to be produced constitutively at maximal or near maximal rates. However, recent reports suggest a role for the *uvrA* and *uvrB* gene products in various inducible repair systems[6,14–17]. Because these studies all relied on indirect measurements of induction, it could never be determined if the results reflected induction of other components of these systems or if the *uvrA* or *uvrB* gene products themselves were regulated. The insertion mutant described here allowed us to monitor expression of the *uvrB* operon directly and provided evidence for regulated synthesis of the *uvrB* gene product.

Our original UV-sensitive isolate, designated CON48, was a derivative of an HfrH strain, CA7092 (see Table 1). To characterize the source of UV-sensitivity in CON48, it was allowed to mate with a multiply-auxotrophic recipient, AB1157. After 30 min the mating pairs were disrupted mechanically and recombinants resistant to both ampicillin and streptomycin were selected. Of 110 Amp^r isolates tested, 74 were found to be Leu$^+$, none were His$^+$, and one was Arg$^+$. Of a further 103 Amp^r colonies chosen, 97 were Gal$^+$. Thus, CON48 seemed to carry the Mu(lac, amp) inserted into a chromosomal segment near the genes for galactose metabolism. The *uvrB* gene is also located in this region and mutants in this gene cause extreme UV-sensitivity[13,21]; thus it seemed reasonable that the Mu(lac, amp) insertion was blocking expression of *uvrB*.

To test this hypothesis and to substantiate the map position of the Mu insertion, the transducing phage Pl$_{clr}$ was grown on CON48. This lysate was used to infect JC3890, a strain deleted for the *uvrB* gene and several adjacent genes, including those for biotin biosynthesis[20,21]. If Mu(lac, amp) was inserted in or near the *uvrB* gene of CON48, we anticipated that all Amp^r transductants would grow in the absence of biotin. This expectation was not fulfilled—most of the Amp^r transductants remained Bio$^-$ while only a small fraction became Bio$^+$. These Bio$^+$ transductants were also Gal$^+$ indicating that they had arisen by a recombination event within the *gal–uvrB* region. One of these isolates was chosen and a temperature-resistant isolate, designated CON481R, which retained its Mu and ampicillin resistance but grew at 42 °C, was selected for further study. The selection of this revertant of the c^{ts} mutation carried on the Mu(lac, amp)[1], was performed so that the cells could be grown at temperatures which produce higher titres of Pl. When Pl lysates were prepared on this host and used to transduce JC3890, 100% of the 132 Amp^r transductants screened also became Bio$^+$ and Gal$^+$. These results suggest that CON48 contains more than one Mu(lac, amp) insertion whereas CON481R carries only one such insertion, that being in or very near the *uvrB* gene.

The *uvrB* region of CON481R was also transferred to AB1157. This produced ampicillin resistant, UV^s transductants. In this recipient only 5 of the 50 Amp^r UV^s transductants screened acquired a Gal$^+$ phenotype. One such transductant was designated CON484 and used along with CON481R in further experiments. Because CON484 had originally been *uvrB*$^+$, any UV-sensitivity it acquired must have resulted from the presence of the insertion mutation.

To establish that the Mu(lac, amp) insertion was in fact causing the cells to be *uvrB*$^-$, two types of experiments were performed. First, the UV-sensitivity of CON484 was compared with that of JC3890 and AB1157. Figure 1 shows that the Mu(lac, amp) insertion produces the same UV-sensitivity as the *uvrB* deletion carried by JC3890. Similar results were obtained with CON481 (data not shown).

Host cell reactivation (HCR) results from the nucleotide excision repair of UV lesions mediated by the *uvrA* and *uvrB* gene products[13]. The ability of CON484 to carry out HCR was compared with that of AB1157 and JC3890. Figure 2 shows that CON484 was as deficient in this process as the *uvrB* deletion strain. These two experiments strongly argue that the Mu(lac, amp) insertion completely blocks *uvrB* expression either because it is inserted within *uvrB*, destroying the integrity of the gene sequence, or because it is inserted within the sequence of DNA coordinately regulated with the *uvrB* gene, but upstream from the gene, thereby totally blocking *uvrB* transcription.

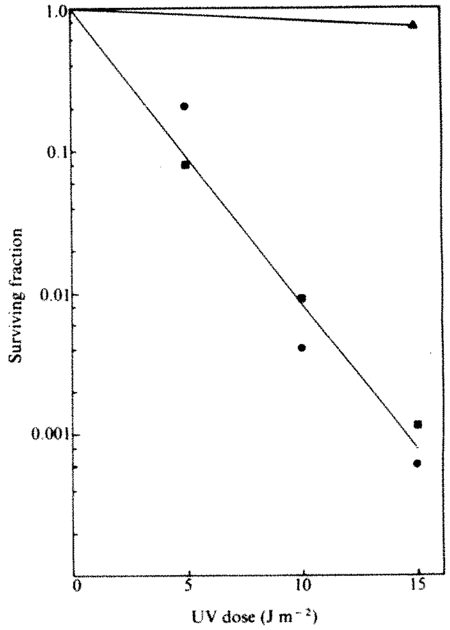


Fig. 1 Cultures (5 ml) of CON484, JC3890 and AB1157 were grown to saturation in L broth[22]. The cells were centrifuged out and resuspended in M9 salts[22] at a density of 2×10^8 ml^{-1}. Samples (1.0 ml) of each culture were irradiated in watchglasses with varied levels of light from a Sylvania G30T8 germicidal lamp, maximum output at 254 nm. Appropriate dilutions were made and the cells plated on LB plates in 2.5 ml of soft agar. Plates were incubated at 30 °C for 36 h before colonies were scored ●, JC3890 (Δ*uvrB*); ▲, AB1157 (*uvrB*$^+$); ■, CON484 (*uvrB*::Mu(lac, amp)). PFU, plaque-forming units.

0028-0836/81/030196—03$01.00

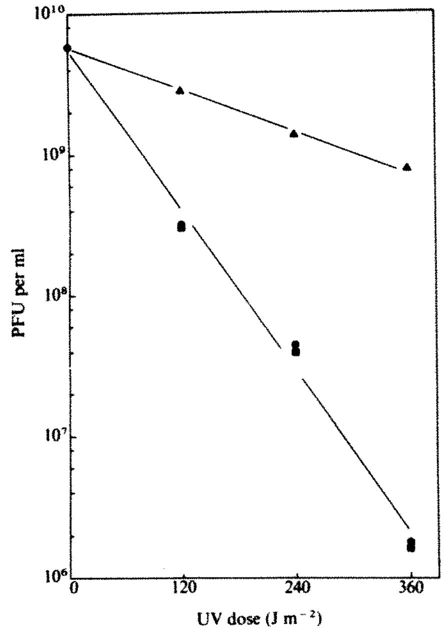


Fig. 2 Cultures of CON484, JC3890 and AB1157 were grown to saturation in T broth containing 0.2% maltose. A stock of λc_1857 in T broth[22] was irradiated in a watchglass with varied levels of 254 nm light. The phage were then appropriately diluted in 0.01 M Tris-HCl (*p*H 7.5) containing 0.01 M $MgSO_4$ and added to 0.1-ml cultures of one of the three strains to be tested. Plaques were scored after 16 h growth on T plates at 37 °C. ●, JC3890 (Δ*uvrB*); ▲, AB1157 (*uvrB*$^+$); ■, CON484 (*uvrB*::Mu(lac, amp)).

To establish finally that the Mu(lac, amp) insertion was located within the *uvrB* operon, strains were constructed in which a plasmid, derived from pBR322 and carrying the chromosomal region which includes the *uvrB* gene[23], was placed in cells carrying the Mu(lac, amp) insertion. This plasmid restored UV resistance to the multiply inserted CON48R as well as strains having only the single *uvrB* insertion. A similar plasmid carrying the same chromosomal segment but with an inactive *uvrB* gene left the insertion strains UV sensitive. We consider this to be very strong evidence that the Mu(lac, amp) insertion is within the *uvrB* operon.

The insertion of Mu(lac, amp) within the *uvrB* operon creates a situation where β-galactosidase is produced in response to stimuli which normally would cause increased synthesis of *uvrB* protein. Figure 3 shows that this operon is highly stimulated by exposure to UV light. During the 3 h following a 10 J m^{-2} exposure, β-galactosidase levels within the cell increase 20-fold (this represents a 5–6-fold increase in β-galactosidase synthesis relative to other proteins). This UV dose produced a near maximum extent of induction. Exposure to other DNA damaging agents, such as *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG), also induces the synthesis of this enzyme but to a lesser extent. Cells without the insertion produce no β-galactosidase and non-irradiated cells continue to produce basal levels of the enyzme throughout the experiment.

The basal level observed from experiment to experiment is rather variable, ranging from 40 to 110 units for strains CON481R and CON484. In contrast to these relatively low basal levels, the original isolate, CON48, has an uninduced level greater than 200 units. Despite this higher level of uninduced β-galactosidase synthesis, its induced level is only slightly higher than that of the strains shown in Fig. 3. This is again consistent with CON48 carrying multiple Mu(lac, amp) insertions—the one in or near *uvrB* is inducible, whereas the other insertion(s) produces 100–200 units of β-galactosidase constitutively.

During the course of this work, a report was published which identified five other gene functions that were induced by UV light[25]. One of these inducible operons was the *uvrA* gene, but no insertion into the *uvrB* gene was reported. All the UV-inducible genes isolated by Kenyon and Walker seem to comprise part of a regulon which is controlled by the *lexA* and *recA* gene products. In an attempt to determine if the *uvrB* gene also belongs to this regulon, *lexA* and *recA* strains carrying the *uvrB*::Mu(lac, amp) insertion were constructed. Both the *lexA*1 and *recA*1 mutations blocked the UV induction of β-galactosidase seen in *lexA*$^+$, *recA*$^+$ insertion strains. Thus, the *uvrB* gene is probably part of the *lexA*, *recA* regulon.

The biological impact of these results is substantial. If the *uvrA*,*uvrB*-mediated excision repair system is controlled by the *lexA*,*recA* regulatory system, and the amount of excision repair capacity increases significantly in response to DNA damage, this could explain the observations of increased survival of UV-ray irradiated phage in cells carrying the *tsl* mutation[14]. It may also be responsible for parts of Weigle-reactivation[16,17], and post-replication repair[6]. If this is the case it could provide a significant simplification to an otherwise bewilderingly complex picture of UV repair.

However, a note of caution must be added. We have presented here evidence consistent with the idea that the *uvrB* gene is

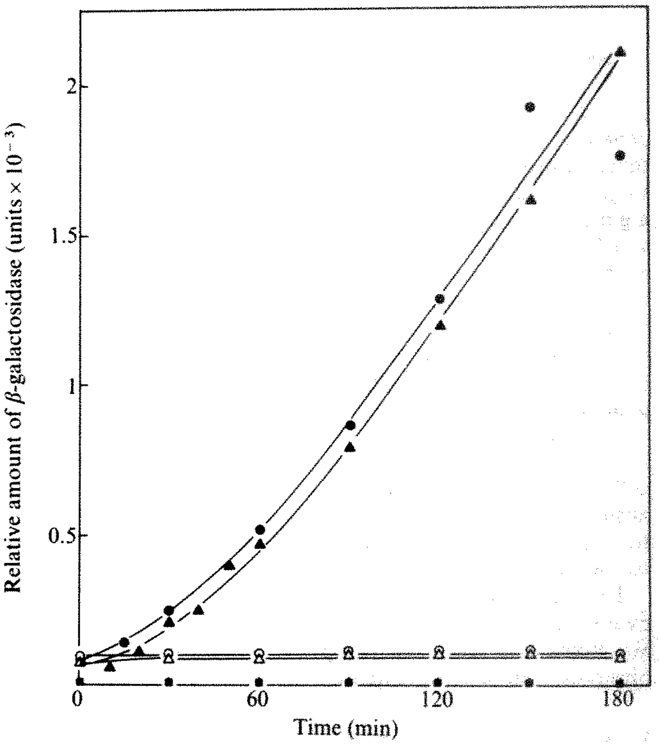


Fig. 3 Strains CON481R and CON484 were grown at 37 °C in M9 salts medium supplemented with 0.2% glucose, 2 μg ml^{-1} thiamine and 40 μg ml^{-1} Arg, Hist, Leu, Pro and Thr. When the cultures reached a cell density of 2–4 × 10^8 ml^{-1}, a 3-ml aliquot was irradiated with 10 J m^{-2} of 254 nm light. A 2-ml sample of these cells was diluted in 8 ml of pre-warmed medium and returned at 37 °C to incubate. The non-irradiated control cells were diluted similarly. Samples (0.5 ml) were withdrawn at various times and added to 0.5 ml of Zcm buffer[24] on ice. The cell density was monitored at 600 nm. Because irradiated cells stopped dividing, the optical density of the initial point was used throughout the outgrowth as an indication of relative cell number. The optical denisty at 600 nm of non-irradiated cultures was monitored at each time point. After completion of sampling, toluene was added to make the cells permeable; the toluene was then evaporated at 37 °C and 0.2 ml of *o*-nitrophenyl-β-D-galactopyranoside (4 mg per ml in 0.1 M sodium phosphate buffer, *p*H 7.0) was added to each sample and the reaction allowed to proceed at 28 °C. The reaction was stopped by addition of 0.5 ml of 1.0 M Na_2CO_3. Relative amount of β-galactosidase is expressed as standard units which have been defined[22] as $[(A_{420}-1.75\,A_{550})/(tvA_{600})]\times 1{,}000$. ○, Non-irradiated CON481R; ●, irradiated CON481R; △, non-irradiated CON484; ▲, irradiated CON484; ■, non-irradiated or irradiated AB1157.

Table 1 Genotypes and origins of strains used

Strain	Genotype	Source
CA7092	HfrH *thi*1 *ara* Δ (*proA*, *B*-*lac*)$_{\mathrm{XIII}}$	Reznikoff*
AB1157	F^- *thr*1 *leu*6 *proA*2 *his*4 *thi*1 *argE*3 *lacY*1 *galK*2 *ara*14 *xyl*5 *mtl*1 *tsx*33 *rpsL*31 *supE*44	Mount†
JC3890	Same as AB1157 plus Δ (*uvrB*-*pgi*)$_{301}$	Kato‡
CON48	HfrH *thi*1 *ara* Δ (*proA*, *B*-*lac*)$_{\mathrm{XIII}}$ multiple inserts of Mu(*lac*, *amp*)cts	Amp^r isolate of CA7092 infected with Mu(*lac*, *amp*)cts§
CON481	Same as AB1157 but carrying the *uvrB* :: *Mu*(*lac*, *amp*)cts from CON48	Amp^r bio^+ gal^+ transductant of JC3890 infected with $\mathrm{Pl}_{\mathrm{clr}}$ grown on CON48§
CON481R	Same as AB1157 but carrying *uvrB* :: *Mu*(*lac*, *amp*)	Temperature-resistant amp^r Mu^r isolate of CON481§
CON484	Same as CON481R	Amp^r Gal^+ transductant of AB1157 infected with $\mathrm{Pl}_{\mathrm{clr}}$ grown on CON481R§

* Ref. 18; † ref. 19; ‡ ref. 20; § this study.

part of an inducible regulon controlled by the *lexA* and *recA* gene products. This conclusion seems to make good biological sense. Unfortunately, it has been based on the analysis of a single insertion mutation. We cannot rule out the possibility that a genetic rearrangement could have occurred during insertion which fused the *uvrB* gene to a promoter other than the one normally used by *uvrB*. If this occurred, we have been analysing regulation at this promoter rather than the one normally used for synthesis of *uvrB* message. Although this is possible, it seems unlikely that a second extraneous promoter fortuitously fused to *uvrB* would be controlled by the *lexA* and *recA* gene products.

Another complication which could affect our conclusion arises from the fact that the Mu(lac, amp) insertion makes the cell UvrB^-. If the *uvrB* gene is in any way autoregulatory, the absence of an active *urvB* gene product could drastically alter the regulation of this operon. Experiments are being carried out to clarify this situation.

This work was supported in part by NCI grant CA 14733 and by the Institute of Materials Science, University of Connecticut. We thank Drs Kacinski, Rupp and Sancar for supplying the plasmids used in complementation analysis of our strains.

Received 14 July; accepted 27 October 1980.

1. **Casadaban, M. J. & Cohen, S. N.** *Proc. natn. Acad. Sci. U.S.A.* **76,** 4530–4533 (1979).
2. **Defais, M., Fauquet, P., Radman, M. & Errera, M.** *Virology* **43,** 495–503 (1971).
3. **Defais, M., Caillet-Fauquet, P., Fox, M. S. & Radman, M.** *Molec. gen. Genet.* **148,** 125–130 (1976).
4. **Sedgwick, S. G.** *Proc. natn. Acad. Sci. U.S.A.* **72,** 2753–2757 (1975).
5. **Youngs, D. A. & Smith, K. C.** *J. Bact.* **125,** 102–110 (1976).
6. **Ganesan, A. K., Seawell, P. C. & Mount, D. W.** *J. Bact.* **135,** 935–942 (1978).
7. **Samson, L. & Cairns, J.** *Nature* **267,** 281–282 (1977).
8. **Jeggo, P., Defais, M., Samson, L. & Schendel, P.** *Molec. gen. Genet.* **157,** 1–9 (1977).
9. **Braun, A. & Grossman, L.** *Proc. natn. Acad. Sci. U.S.A.* **71,** 1838–1842 (1974).
10. **Braun, A. G., Radman, M. & Grossman, L.** *Biochemistry* **15,** 4116–4120 (1976).
11. **Seeberg, E.** *Proc. natn. Acad. Sci. U.S.A.* **75,** 2569–2573 (1978).
12. **Witkin, E. M.** *Bact. Rev.* **40,** 869–907 (1976).
13. **Howard-Flanders, P. & Boyce, R. P.** *Radiat. Res. Suppl.* **6,** 156–184 (1966).
14. **Mount, D. W., Kosel, C. K. & Walker, A.** *Molec. gen. Genet.* **146,** 37–41 (1976).
15. **Clark, A. J. & Volkert, M. R.** in *DNA Repair Mechanisms* (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 57–72 (Academic, New York, 1978).
16. **Rothman, R. H., Margossian, L. J. & Clark, A. J.** *Molec. gen. Genet.* **169,** 279–287 (1979).
17. **Walker, G. C. & Dobson, P. P.** *Molec. gen. Genet.* **172,** 17–24 (1979).
18. **Cuzin, F. & Jacob, F.** *C.r. hebd. Séanc. Acad. Sci., Paris* **258,** 1350–1352 (1964).
19. **Howard-Flanders, P., Boyce, R. P. & Theriot, L.** *Genetics* **53,** 1119–1136 (1966).
20. **Kato, T., Rothman, R. H. & Clark, A. J.** *Genetics* **87,** 1–18 (1977).
21. **Bachmann, B. J. & Low, K. B.** *Microbiol. Rev.* **44,** 1–56 (1980).
22. **Miller, J. H.** in *Experiments in Molecular Genetics* (Cold Spring Harbor Press, New York, 1972).
23. **Rupp, W. D., Sancar, A., Kennedy, W. J., Ayers, J. & Griswold, J.** in *DNA Repair Mechanisms* (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 229–235 (Academic, New York, 1978).
24. **Elespuru, R. K. & Yarmolinsky, M. B.** *Envir. Mutat.* **1,** 65–78 (1979).
25. **Kenyon, C. J. & Walker, G. C.** *Proc. natn. Acad. Sci. U.S.A.* **77,** 2819–2823 (1980).

Chromatin fine structure of active and repressed genes

Abraham Levy & Markus Noll

Department of Cell Biology, Biocenter of the University,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Study of the structural organization of chromatin during transcription and replication may reveal important aspects of these processes. At the lowest level of organization, chromatin consists of a repeating subunit, the nucleosome (for reviews see refs 1–3). Electron microscopy indicates that the nucleosomes are arranged helically[4–6] or form discrete superbeads[7], generating the familiar 250 Å–300-Å fibre[8]. It has been suggested that this fibre is further folded into loops containing up to several hundred nucleosomes[9,10]. Despite extensive study, the significance and fate of these nucleosomes remain obscure. We have used here micrococcal nuclease digestion to compare the structures of actively transcribing and inert chromatin of the genes coding for the major heat-shock protein of *Drosophila melanogaster*. The repressed *hsp 70* genes were considerably more resistant to cleavage by micrococcal nuclease than their flanking regions and the bulk of chromatin. The active genes, previously shown to be more sensitive than the repressed genes[11–13], are also more susceptible to the nuclease than their 3′-flanking regions and bulk chromatin.

The heat-shock genes of *D. melanogaster* provide an excellent system for studying different aspects of gene structure and regulation (reviewed in ref. 14). These genes are induced by transferring *Drosophila* tissues or cell cultures from their normal temperature at 25 °C to elevated temperatures above 30 °C. The major protein synthesized at 37 °C—*hsp 70*—has a molecular weight of 70,000. Six copies of the gene coding for this protein exist in the haploid genome of Kc cells[15]; each has been extensively characterized by cloning and whole-genome Southern analysis[15] (M.-E. Mirault, personal communication).

Nuclei of a *Drosophila* tissue-culture cell line were digested with micrococcal nuclease before or after heat shock. The DNA fragments obtained were separated according to size by agarose gel electrophoresis, denatured, transferred and bound covalently to diazobenzyloxymethyl (DBM)-paper[16]. By hybridization with radioactively labelled probes, the DNA sequences of interest were analysed selectively. Three different probes were used (Fig. 1)—the *Sal*–*Sal* fragment of 56H8 containing almost the entire coding region yet no flanking region, the *Sal*–*Bgl*II fragment of 56H8 adjacent to the 3′ end of the coding region and the *Bgl*II–*Bgl*II fragment to the left of the coding region in the genomic clone 132E3. The last probe contains moderately repetitive sequences[17] and was found to exhibit the same distribution of the nucleosomal repeat pattern (Fig. 2*C*) as bulk chromatin [revealed by staining the gel with ethidium bromide (Fig. 2*A*)] and hence is designated 'bulk'. Whenever the patterns revealed by these different probes are compared, they have been obtained by hybridizations with the same DBM-paper after release of the previous probe by denaturation.

The micrococcal nuclease pattern of the coding region (Fig. 2*B*) differs from that of bulk chromatin (Fig. 2*A*, *C*) in non-heat-shock conditions. The higher ratios of multimers to monomer in the lanes of the repressed gene (Fig. 2*B*) compared with the corresponding lanes representing bulk chromatin (control lanes in Fig. 2*A*, *C*) indicate that at least part of the coding region of the repressed gene is more resistant to micrococcal nuclease than bulk chromatin. This conclusion is corroborated by the observation that the average size of nick-translated DNA (bulk) obtained after nuclease digestion of

0028-0836/81/030198—06$01.00

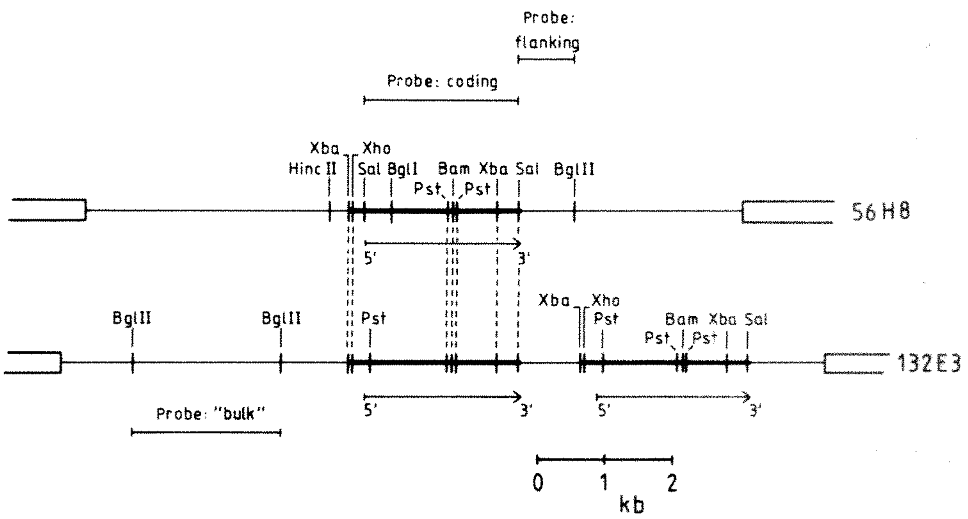


Fig. 1 Restriction map of *hsp 70* genes. Restriction maps of the *D. melanogaster* sequences in the hybrid plasmids 56H8 and 132E3 (ref. 26) have been published[27]. Only a few selected restriction sites are indicated here; the distances have been slightly corrected by calibration with additional low-molecular-weight markers (*Hae*III digest of PM2 DNA[28]). Sequences complementary to the *hsp 70* mRNA and the direction of transcription of these genes are indicated by arrows[18]. The protected domain of the repressed *hsp70* genes is represented by a thick bar comprising the coding region and most of the z_{nc} region[18]. The probes for the coding (*Sal–Sal*) and the flanking region (*Sal–Bgl*II), and the probe revealing a pattern similar to bulk chromatin (*Bgl*II–*Bgl*II) are shown above and below the maps. The names of the restriction enzymes have been abbreviated as follows: *Bam*, *Bam*HI; *Pst*, *Pst*I; *Sal*, *Sal*I; *Xba*, *Xba*I; *Xho*, *Xho*I. kb, Kilobases.

non-heat-shocked nuclei is much smaller than that of the pattern revealed by hybridization of the same DNA with the coding region. (Compare lanes labelled 'nt' and 'c' of 6% acid-soluble DNA in Fig. 3. Considering that the label incorporated into nick-translated DNA is proportional to its mass whereas the label of the hybridized DNA is proportional to the number of DNA molecules, the observed difference is even greater than apparent from a mere comparison of the patterns of the autoradiogram.)

After mild digestions, the DNA of the repressed gene hybridizing to the coding region exhibits a relatively narrow size distribution with an average of about 2.5 kilobases (control lane of *a* in Fig. 2*B*). In a more extensive digestion, three sharp bands at 2.52, 2.34 and 2.16 ± 0.05 kilobases appear above a background (control lane of *b* in Fig. 2*B*). These bands are more clearly visible when the background is reduced by a shorter exposure (lane *b* at far right in Fig. 2*B*). Therefore, a region of 2.5 kilobases, larger than the coding region[18] but which must contain at least part of the coding region, is more resistant to micrococcal nuclease digestion than its flanking regions when the gene is not expressed. To determine the right boundary of this protected domain, the *Sal–Bgl*II probe specific for the 3'-flanking region of one of the genes (Fig. 1) was used. As evident from Fig. 2*D*, the patterns resemble those of the bulk DNA (Fig. 2*A*, *C*) more closely than those of the coding region (Fig. 2*B*). The three bands between 2 and 2.5 kilobases are barely visible (control lane of *b* in Fig. 2*D*), and digestion is slightly inhibited in the repressed gene (ratio of monomer to oligomer DNA in lanes of control in Fig. 2*D*) compared with the bulk (Fig. 2*C*). This suggests that the protected domain and the *Sal–Bgl*II region overlap only slightly. In other words, the protected region of the repressed gene ends at a site close to the 3' end of the mRNA coding region[18] (Fig. 1).

From the size of the protected domain of about 2.5 kilobases, we predict that its left end reaches beyond the 5' end of the mRNA coding region and maps close to the left boundary of the z_{nc} region[18]. A more precise localization of the 5' limit by direct mapping is complicated by the presence of repetitive sequences in this region.

No bands are ever observed between 2 and 2.5 kilobases in heat-shock conditions, even when digestion is very mild ($<0.2\%$ acid-soluble DNA in Fig. 3), so that the DNA containing the coding region exhibits about the same average size as the protected domain of the repressed gene. (The size of the bulk DNA in such digests is on average 15 kilobases, as shown in the two lanes on the left in Fig. 3.)

During more extensive digestions, beyond 7% acid-soluble DNA, the protected domain of the inactive *hsp 70* genes is eventually degraded to smaller multiples of the nucleosomal repeat which are resolved significantly better than in bulk chromatin (control lane of *c* in Fig. 2*B*, and 11% acid-soluble DNA in Fig. 3). The largest fragments observed correspond to 12, 13 and 14 nucleosomes and co-migrate with the three bands of 2.16, 2.34 and 2.52 kilobases visible after milder digestions, whereas nucleosomal repeats larger than 10 are not resolved in bulk chromatin. It follows that the average spacing of nucleosomes is more uniform in the inactive gene than in bulk chromatin. Calibration with markers of known lengths shows that the nucleosomal repeat of the inactive gene is 180 ± 4 base pairs (Fig. 3).

The fragment sizes resulting from digestion of active genes are not distributed randomly, yet differ from those obtained after micrococcal nuclease digestion of bulk chromatin or of inactive genes. This results in bands ('hs' lanes of *b* and *c* in Figs 2*B* and 4) at positions between those of the familiar nucleosomal repeat. These bands are clearer in Fig. 4 where the DNA has been analysed in denaturing conditions. As the bands appear at roughly equivalent positions and intensities in denaturing (Fig. 4) and non-denaturing gels (Fig. 2*B*), the DNA fragments produced are largely double-stranded. These bands do not originate from the presence of preferential cleavage sites in the free DNA because digestion of free DNA of a mixture of 56H8 and 132E3 does not produce such bands after hybridization to the same *Sal–Sal* probe. Nor are they oligonucleosomes of which the DNA has been degraded from the ends because no such monosomes are observed. Moreover, in contrast to bulk chromatin, these structures do not show any metastable intermediates at any level of digestion but are readily degraded to very short pieces. This is demonstrated in Fig. 4 in which DNA fragments as short as 57 nucleotides have been transferred efficiently (shortest marker indicated on the left), yet very little hybridization is observed below 160 bases. The conditions of hybridization used permit detection of hybrids as short as 40 base pairs[19].

The observed changes in chromatin structure depend completely on the level of transcription. After returning heat-shocked cells to the normal temperature (25 °C), there is a gradual disappearance of the transcription-specific structures

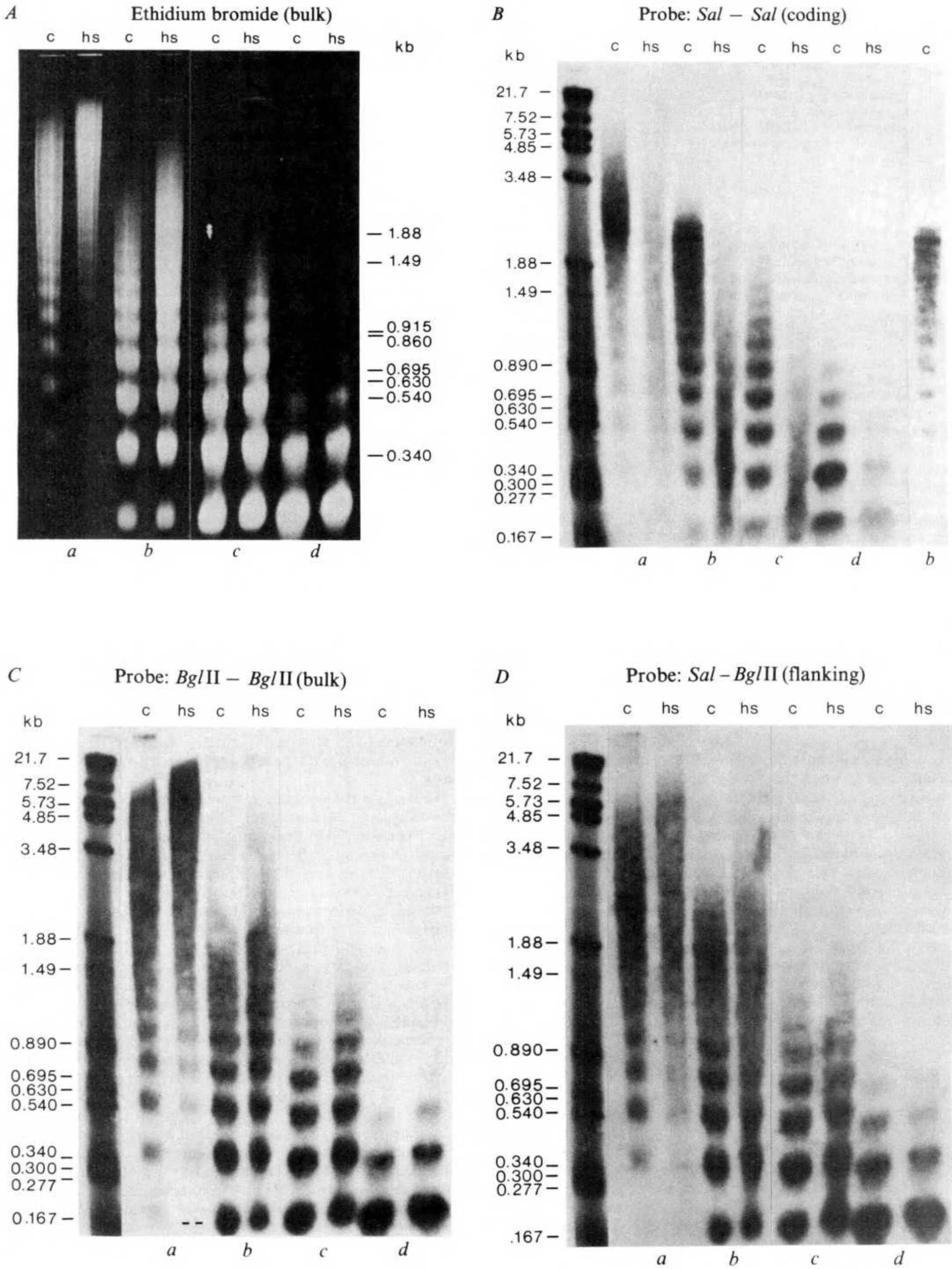

A
Ethidium bromide (bulk)
c hs c hs c hs c hs
kb
1.88
1.49
0.915
0.860
0.695
0.630
0.540
0.340
a b c d
B
Probe: *Sal* – *Sal* (coding)
kb
c hs c hs c hs c hs c
21.7
7.52
5.73
4.85
3.48
1.88
1.49
0.890
0.695
0.630
0.540
0.340
0.300
0.277
0.167
a b c d b
C
Probe: *Bgl*II – *Bgl*II (bulk)
kb
c hs c hs c hs c hs
21.7
7.52
5.73
4.85
3.48
1.88
1.49
0.890
0.695
0.630
0.540
0.340
0.300
0.277
0.167
a b c d
D
Probe: *Sal* – *Bgl*II (flanking)
kb
c hs c hs c hs c hs
21.7
7.52
5.73
4.85
3.48
1.88
1.49
0.890
0.695
0.630
0.540
0.340
0.300
0.277
.167
a b c d

Fig. 2 Chromatin structure of *hsp 70* genes: comparison with structure of bulk and flanking regions. Digests of control (labelled 'c') and heat-shocked nuclei ('hs') were compared pairwise in bulk chromatin (*A*), in the coding region of the *hsp 70* gene (*B*), in a region containing a middle-repetitive sequence (*C*) and in the 3′-flanking region of the *hsp 70* gene (*D*). Four levels of digestion are shown in each panel corresponding to about 2.3% (*a*), 7% (*b*), 12% (*c*) and 18% (*d*) acid-soluble DNA. For preparation of nuclei, *Drosophila* Kc cells (adapted to growth in low serum) were grown in suspension at 25 °C in Echalier's medium[29] supplemented with 1% fetal calf serum (Gibco). Cells heat-shocked for 35 min at 36 °C, or control cells not subjected to heat shock, were collected in mid-log phase at a density of about 8×10^6 cells ml^{-1}. The cells were quickly chilled by pouring them on to half a volume of frozen balanced salt solution (BSS), and all subsequent steps were carried out on ice or at 4 °C. The cells were pelleted, washed twice in BSS and resuspended in Mg^{2+}-buffer (0.33 M sucrose, 20 mM HEPES, *p*H 7.5, 80 mM NaCl, 1 mM $MgCl_2$) at 4×10^8 cells ml^{-1}. To lyse the cells, the cellular suspension was made 0.2% in Triton X–100, incubated for 15 min on ice and passed four times through a hypodermic needle (G-18). The nuclei were washed (1,400*g* for 10 min) two or three times in Mg^{2+}-buffer and finally resuspended at a concentration of 1.2×10^9 nuclei ml^{-1}. For digestion, $CaCl_2$ to 1 mM was added to the nuclear suspension, and 0.5-ml aliquots were incubated for 2 min at 37 °C with 20 (*a*), 60 (*b*), 180 (*c*) and 360 units (*d*) of micrococcal nuclease (Worthington). The DNA was extracted and treated with a mixture of RNase A and T_1 RNase. Samples containing 5 μg (*A*) or 50 μg of DNA (*B–D*) were subjected to electrophoresis in two 1.7% agarose slab gels in 50 mM KNa_2PO_4, *p*H 6.5, 4 mM NaAc, 0.5 mM EDTA. One gel was stained with ethidium bromide (*A*), while the DNA from the other gel was transferred to DBM-paper and analysed by hybridization with various probes (Fig. 1); *Sal–Sal* of 56H8 (obtained as subclone *Sal*-O; *B*), *Bgl*II–*Bgl*II of 132E3 (*C*) and *Sal–Bgl*II of 56H8 (*D*). Before transfer, the DNA was denatured by washing the gel twice for 15 min in 0.5 M NaOH, neutralized in 0.5 M Na-phosphate, *p*H 5.5, for 15 min, and washed twice for 10 min in 11.5 mM Na-phosphate-citrate, *p*H 4.0, first at room temperature, then at 4 °C. For preparation of DBM-paper, a published procedure[16] was modified to achieve an efficiency of DNA transfer of at least 40% over a size range of 40 bases to 21 kilobases. Using this modification, we can transfer at least twice as much DNA as was used in the experiments reported here before the binding capacity of the DBM-paper is reached[19]. Hybridization with the nick-translated[30], ^{32}P-labelled probe was carried out in 0.02% each of Ficoll, polyvinylpyrrolidone and bovine albumin[31], 5 × SSC, 1 mM EDTA, 0.1% SDS, 2 mg ml^{-1} salmon sperm DNA. After hybridization, the paper was washed four times for 1 h at 65 °C in 2 × SSC, 0.5% SDS, 17 mM Na-phosphate, *p*H 7.5, 0.05% $Na_4P_2O_7$, 1 mM EDTA and finally briefly in water at room temperature. For autoradiography, preflashed[32] Kodak XR-5 film was used with a Kyokko intensifying screen. All three autoradiograms have been obtained from one transfer of DNA to DBM-paper. After autoradiography, the hybridized probe was removed by four 30-min washes in 99% formamide at 80 °C, the DBM-paper was washed in water and the remaining ^{32}P-radioactivity (<3%) further reduced during 3–4 half lives to less than 0.3%. For calibration, the far-left lanes of each autoradiogram show *Hae*III fragments of PM2 DNA[28] and *Eco*RI fragments of λ DNA[33] labelled at their 3′ ends by 'Klenow' DNA polymerase I (Boehringer Mannheim). The lane on the far right in (*B*) represents a shorter exposure of the fourth lane from the left in (*B*). kb, kilobases.

(not shown). In some experiments (not shown), after micrococcal nuclease cleavage in the mRNA coding region of the active gene, we observed a smeared hybridization pattern initially which then gave way to a pattern reflecting the nucleosomal repeat. We could clearly correlate this smearing effect with a suboptimal heat-shock response. It thus seems crucial that transcription-specific structures be observed in heavily transcribed genes.

If the digestions after heat shock are relatively extensive ('hs' lanes of *d* in Figs 2*B*, 4), most of the hybridized DNA disappears, and only faint bands at the monomer and dimer positions of the inactive gene are visible. The residual monomer and dimer DNA can be explained if a few cells do not respond to heat shock or if a minor fraction of active genes contains small regions of a structure similar to that of inactive genes. Inactivity of one or more of the six *hsp 70* genes in all cells after heat shock is unlikely because, as the milder digestions show, the quantitative differences between the active and repressed genes exceed a ratio of 6 : 1 (compare, for example, the region between 2 and 2.5 kilobases in the two lanes of *b* in Fig. 2*B*).

As even bulk chromatin was more sensitive to micrococcal nuclease than the repressed gene, it was of interest to compare the sensitivity of the active gene with that of bulk chromatin. Comparison of the corresponding lanes of digests obtained after heat shock in Fig. 2*B* with *C* shows a smaller average DNA size and a reduced amount of hybridization in the former. Thus the DNA of active genes is more sensitive than the DNA of bulk chromatin.

We also compared the initial rates of digestion of the active and repressed genes. The average DNA size of the active gene after digestion to less than 0.2% acid-solubility is about the same as that of the repressed gene after digestion to 6% acid-solubility for which 30 times more enzyme is required (Fig. 3). Thus the active gene is degraded 30 times more rapidly.

In addition, we conclude from a comparison of the amounts of hybridizable material in denaturing conditions in control and heat-shock DNA (for example ~1.9 kilobases in lanes of *b* or ~1 kilobase in lanes of *c* in Fig. 4) that both DNA strands of the active gene are more sensitive to micrococcal nuclease than those of the repressed gene.

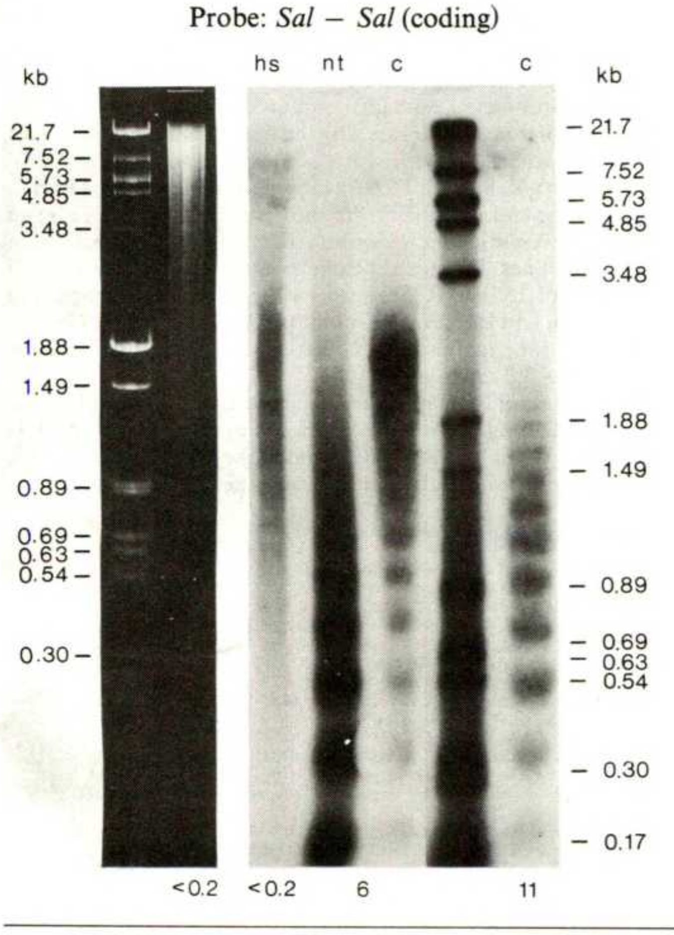


Fig. 3 Nucleosome spacing in repressed *hsp 70* genes and relative sensitivities of active and repressed *hsp 70* genes to micrococcal nuclease. On the left of the autoradiogram, the relative sensitivities to micrococcal nuclease of the active *hsp 70* gene ('hs'), of bulk chromatin ('nt') and of the repressed *hsp 70* gene ('c'; 6% acid-soluble DNA) are compared. At the far right, the uniform nucleosomal repeat of the repressed gene is shown ('c'; 11% acid-soluble DNA). Nuclei of heat-shocked ('hs') and control cells ('c') were prepared as in Fig. 2 except that the Mg^{2+}-buffer was replaced by spermine-spermidine buffer[34] (0.33 M sucrose, 15 mM Tris-HCl, *p*H 7.5, 60 mM KCl, 15 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine) and the concentration of Triton X-100 was raised to 0.4%. Digestion of heat-shocked nuclei (1.2×10^9 ml^{-1}) with 4 U ml^{-1} of micrococcal nuclease and digestion of control nuclei with 120 or 360 U ml^{-1} rendered <0.2%, 6% and 11% of the DNA acid-soluble, as indicated below the corresponding lanes. The DNA was extracted, and 50-μg samples were subjected to electrophoresis in 1.7% agarose gels, transferred to DBM-paper and hybridized to labelled *Sal*-O DNA as described in Fig. 2. The 'nt' lane shows nick-translated DNA of the digest generating 6% acid-soluble DNA. The second lane from the right contains the same markers as in Fig. 2. The two lanes on the left show the same marker DNA and DNA of the digest of less than 0.2% acid-solubility after staining with ethidium bromide.

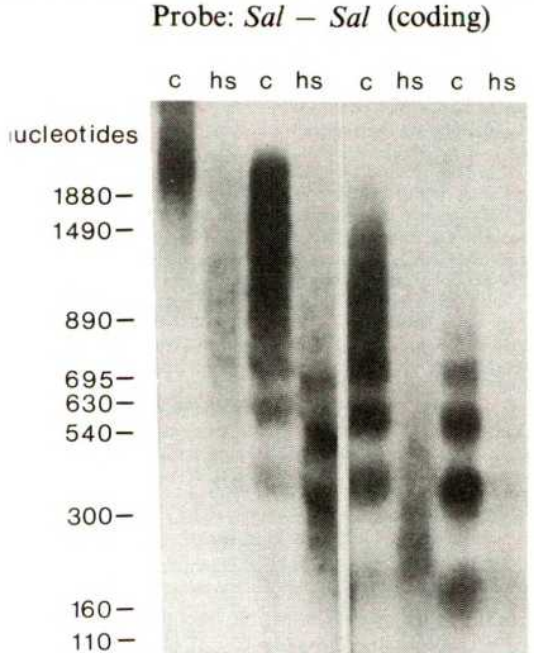


Fig. 4 Analysis of DNA fragments of the *hsp 70* gene in denaturing conditions. DNA samples (50 μg) of the same digests of control ('c') and heat-shocked nuclei ('hs') shown in Fig. 2 were subjected to electrophoresis in an alkaline 1.7% agarose gel[35]. After electrophoresis, the DNA was neutralized, transferred to DBM-paper, hybridized to *Sal*-O DNA and autoradiographed as described in Fig. 2 legend. On the left, the positions of *Hae*III fragments of PM2 DNA, used for calibration, are indicated.

As genes may be part of large loops of domains[9,10,20,21], it was interesting to examine whether the sensitive region of the active state extends beyond the coding region to include such larger domains. A comparison of the hybridization pattern of the 3'-flanking region (Fig. 2*D*) with that of the coding region (Fig. 2*B*) shows that the enhanced sensitivity does not extend much beyond the 3' end of at least one of the *hsp 70* genes (a similar study with the 5'-flanking region is complicated by the presence of repetitive sequences). The pattern obtained by hybridization to the 3'-flanking-region probe (Fig. 2*D*) is very similar to that of the bulk (Fig. 2*A*, *C*). Thus, for a first approximation, the region downstream from the right *Sal* site (Fig. 1) is organized similarly to bulk chromatin in heat-shock conditions. However, the patterns revealed by these two probes do differ slightly. First, the heat-shock lanes of *b* and *c* in Fig. 2*D* show more hybridization in the regions between monomer and dimer and between dimer and trimer DNA than the corresponding non-heat-shock controls or than the same lanes representing bulk chromatin (Fig. 2*A*, *C*). This suggests that the *Sal*–*Bgl*II probe overlapped slightly with the chromatin region containing the transcription-specific structure (Fig. 2*B*). Second, there is less hybridization after extensive digestion of the heat-shock-treated nuclei ('hs' lanes of *d*) in the flanking region (Fig. 2*D*) compared with bulk DNA (Fig. 2*C*) when normalized to the control. We infer that the sensitive portion of the gene ends close to the right-hand *Sal* site. As the 3' end of the mRNA coding region also maps close to this site[16], the 3' end of the mRNA coding region and the right boundary of the nuclease-sensitive chromatin region must be very close to each other.

Altered sensitivity to nuclease of actively transcribed genes may result from several structural changes: (1) changes in the higher structural orders, (2) changes in the structure of the nucleosomes, and (3) the absence of histones from transcribed DNA. Therefore, earlier suggestions made solely on the basis of DNase sensitivity that active genes contain structurally modified nucleosomes are invalid[22,23]. More detailed information with respect to the structure of active genes may be gained by analysing the size distribution of the DNA fragments produced by the action of DNase, and this study has revealed transcription-specific structures not previously observed. Micrococcal nuclease was found to discriminate between a gene, regardless of its state, and the rest of the chromatin. The repressed genes were protected and the active genes were highly sensitive compared with bulk chromatin or the neighbouring non-transcribed region. The levels of sensitivity of the gene are reversible and depend on its transcriptional activity. Hence, models of gene repression need not be based solely on the interaction of regulatory proteins with the 5' end. Similarly, the association of bulk-type nucleosomes with the gene seems, *in* itself, insufficient to keep the gene in the inactive form. Some mechanism seems to exist which, on repression, alters the chromatin structure of the gene so as to reduce its accessibility (and that of a region of about 1 nucleosome at its 5' end) to the nuclease considerably below that of the flanking, non-transcribed DNA sequence.

The mechanism of protection of the repressed genes could be explained on at least two structural levels. Each linker region joining adjacent nucleosomes could be modified in such a way as to reduce its sensitivity to micrococcal nuclease. The known sequence of the *hsp 70* gene[24,25] rules out the trivial explanation that all linker regions are less sensitive to micrococcal nuclease because of a high G+C content. In a more attractive model, all the linker regions of the entire domain become collectively resistant to the nuclease by some change in higher order of chromatin structure. The transition points between the protected domain of the repressed state and the flanking, bulk-like structure might then be sequences of regulatory significance. Specific nonhistone protein(s) interacting with these sequences may associate with each other to bring the two ends of this chromatin segment into tight proximity, producing a loop containing 14 nucleosomes. Such a loop may form a supercoiled, more compact and less accessible form of chromatin. In this model, activation of a gene would require at least two steps. First, the protected domain is unfolded, and only then does initiation of transcription start by the modification or removal of the nucleosomes.

We thank P. Schedl, S. Artavanis-Tsakonas, M.-E. Mirault, W. Gehring and A. Tissières for the hybrid plasmids 132E3 and 56H8, M. Goldschmidt-Clermont for the *Sal*-O subclone of 56H8, R. Jack for the subclone containing the *Bgl*II–*Bgl*II fragment of 132E3, R. Steward for the Kc cell line adapted to low serum, and E. Frei, M. Goldberg, R. White and H. Noll for criticism of the manuscript. This work was supported by Swiss NSF grants 3.719.76 and 3.466.79.

The results of this paper have been presented previously at the Annual Meeting of Swiss Societies of Experimental Biology in Basel, on March 13–14, 1980[36] and at the Gordon Conference on Chromatin and Nuclear Proteins (June 15–20, 1980).

Received 17 September; accepted 27 November 1980.

1. Kornberg, R. D. *A. Rev. Biochem.* **46,** 931–954 (1977).
2. Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42,** 1209–1234 (1978).
3. McGhee, J. D. & Felsenfeld, G. *A. Rev. Biochem.* **49,** 1115–1156 (1980).
4. Filip, D. A., Gilly, C. & Mouriquand, C. *Humangentik* **30,** 155–165 (1975).
5. Finch, J. T. & Klug, A. *Proc. natn. Acad. Sci. U.S.A.* **73,** 1897–1901 (1976).
6. Ris, H. & Korenberg, J. in *Cell Biology* Vol. 2 (eds Prescott, D. M. & Goldstein, L.) 267–361 (Academic, New York, 1979).
7. Hozier, J., Renz, M. & Nehls, P. *Chromosoma* **62,** 301–317 (1977).
8. Ris, H. & Kubai, D. F. *A. Rev. Genet.* **4,** 263–294 (1970).
9. Benyajati, C. & Worcel, A. *Cell* **9,** 393–407 (1976).
10. Paulson, J. R. & Laemmli, U. K. *Cell* **12,** 817–828 (1977).
11. Panet, A. & Cedar, H. *Cell* **11,** 933–940 (1977).
12. Bellard, M., Gannon, F. & Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42,** 779–791 (1978).
13. Wu, C., Wong, Y.-C. & Elgin, S. R. C. *Cell* **16,** 807–814 (1979).
14. Ashburner, M. & Bonner, J. J. *Cell* **17,** 241–254 (1979).
15. Mirault, M.-E., Goldschmidt-Clermont, M., Artavanis-Tasakonas, S. & Schedl, P. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5254–5258 (1979).

16. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
17. Lis, J. T., Prestidge, L. & Hogness, D. S. *Cell* **14,** 901–919 (1978).
18. Artavanis-Tsakonas, S., Schedl, P., Mirault, M.-E., Moran, L. & Lis, J. *Cell* **17,** 9–18 (1979).
19. Levy, A., Frei, E. & Noll, M. *Gene* **11,** 283–290 (1980).
20. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. *Cell* **3,** 315–325 (1974).
21. Igó-Kemenes, T. & Zachau, H. G. *Cold Spring Harb. Symp. quant. Biol.* **42,** 109–118 (1978).
22. Weintraub, H. & Groudine, M. *Science* **193,** 848–856 (1976).
23. Garel, A. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3966–3970 (1976).
24. Török, I. & Karch, F. *Nucleic Acids Res.* **8,** 3105–3123 (1980).
25. Ingolia, T. D., Craig, E. A. & McCarthy, B. J. *Cell* **21,** 669–679 (1980).
26. Schedl, P. *et al. Cell* **14,** 921–929 (1978).
27. Moran, L. *et al. Cell* **17,** 1–8 (1979).
28. Noll, M., Zimmer, S., Engel, A. & Dubochet, J. *Nucleic Acids Res.* **8,** 21–42 (1980).
29. Echalier, G. & Ohanessian A. *In Vitro* **6,** 162–172 (1970).
30. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
31. Denhardt, D. T. *Biochem. biophys. Res. Commun.* **23,** 641–646 (1966).
32. Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82,** 314–316 (1977).
33. Philippsen, P., Kramer, R. A. & Davis, R. W. *J. molec. Biol.* **123,** 371–386 (1978).
34. Hewish, D. R. & Burgoyne, L. A. *Biochem. biophys. Res. Commun.* **52,** 504–510 (1973).
35. McDonell, M. W., Simon, M. N. & Studier, F. W. *J. molec. Biol.* **110,** 119–146 (1977).
36. Levy, A. & Noll, M. *Experientia* **36,** 750 (1980).

Theory agrees with experimental thermal denaturation of short DNA restriction fragments

Albert S. Benight*, Roger M. Wartell*† & Dennis K. Howell*

Schools of Physics* and Biology†, Georgia Institute of Technology, Atlanta, Georgia 30332

Experimental melting transitions of several natural DNAs of known nucleotide sequences have recently been obtained. The differential melting curves of these DNAs—ΦX174 DNA[1–3], fd DNA[4] and SV40 DNA[5]—all show distinctive sets of peaks or fine structure. Theoretical melting curves calculated from the sequences and a few *a priori* parameters have not accurately predicted the experimental transitions[2,6,7]. Although calculated fine structure resembled experimental curves in some cases, the characteristic features of a DNA's differential melting curve could not generally be produced. Azbel[8,9] and Gabbarro-Arpa *et al.*[5] have recently obtained good agreement between calculated and experimental curves using a different theoretical approach—only ground-state configurations of DNA were considered for temperatures inside the transition region. Their results suggest that the basic model of DNA melting, common to all theoretical approaches, is accurate. We have used here an exact theoretical approach to calculate melting curves of four DNA restriction fragments of 95–301 base pairs containing the lactose promoter region (Fig. 1). Theoretical curves agree very well with the experimental transitions published by Hardies *et al.*[10] and obtained in this laboratory.

Figure 1 shows the four fragments relative to the *lac* promoter–operator region. The base sequence of this region was determined by Dickson *et al.*[11]. Figure 2 shows the comparison between theory and experiment for the *lac* fragments in 0.1 M Na^+. Theoretical curves predict the melting temperatures, T_ms, peak heights and widths at half height for the four fragments. Although short fragments do not have as much detail in their melting curves as long DNAs, the fine structure observed in Fig. 2 is predicted.

Theoretical curves were generated using the loop entropy model of the DNA helix–coil transition[12–14]. Poland's algorithm including the effects of final dissociation of duplex DNA to single strands was used[15]. Parameter values for T_{AT} and T_{GC}, the average melting temperatures of A·T and G·C base pairs, σ, the nearest-neighbour cooperativity factor, and other thermodynamic parameters are given in Fig. 2 legend. For the short DNAs examined, the melting of internal regions and the corresponding loop entropy contribution to the theory are not significant. Proper account of dissociation of duplex to single strands[16] is essential for accurate agreement with experimental results.

The total fraction of hydrogen-bonded base pairs, θ_{net}, depends on the DNA's external, as well as internal, degrees of freedom. θ_{net} may be written as $\theta_{net} = \theta_{int}\theta_{ext}$, where θ_{int} is the fraction of bonded base pairs among the duplex strands and θ_{ext} is the fraction of strands associated as a duplex with one or more intact base pairs. The equilibrium dissociation of the duplex to single strands may be represented by the reaction:

$$C_2 \rightleftarrows 2C_1 \qquad (1)$$

where C_2 is a duplex and C_1 is a single strand. It has been shown[13,16] that θ_{ext} depends on Z_{int}, the internal molecular partition function of the DNA, C_0, the total strand concentration, and β, the ratio of partition functions for the external degrees of freedom of duplex and single strands.

The mass dependence of rotational and translational partition functions of polyatomic molecules suggests that β varies with DNA length in base pairs, N, as

$$\beta = \frac{Z_{ext}(C_2)}{[Z_{ext}(C_1)]^2} = K\frac{N^{\alpha_2}}{N^{2\alpha_1}} = KN^{\alpha} \qquad (2)$$

where $\alpha = \alpha_2 - 2\alpha_1$ and K is a constant which includes all terms independent of N[13,17]. α_1 and α_2 combine the N dependence of the translational degrees of freedom, $N^{3/2}$, as well as the N dependence of the rotational degrees of freedom. The latter vary according to the molecule's principal moments of inertia, for example, for a rod $N^{7/2}$, for a sphere $N^{5/2}$. Our first attempts to fit theoretical curves to experiment modelled equation (1) as a rod dissociating to two spheres. This assumption yields $\alpha = -3$ for equation (2). A constant K value was selected to provide a

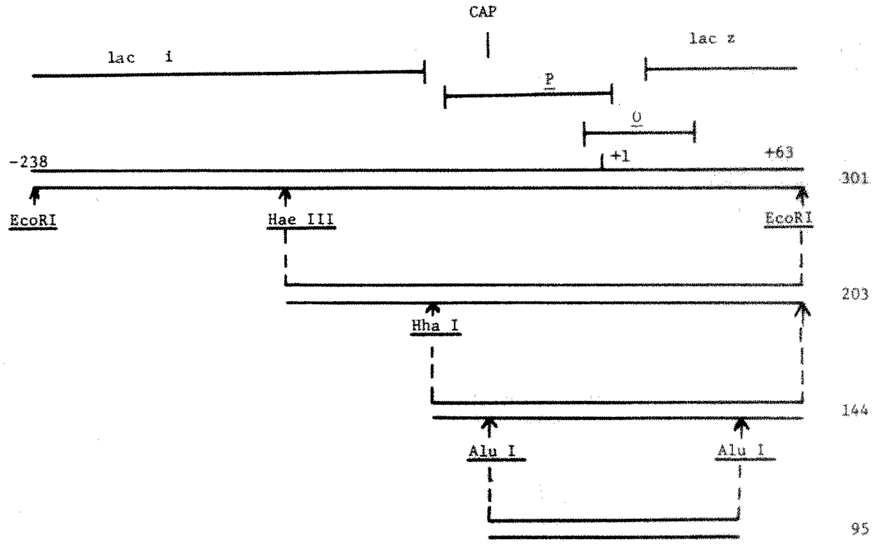


Fig. 1 The 95-, 144-, 203- and 301-base pair restriction fragments containing part of the lactose operon transcription initiation region. Each fragment is fully embodied within the next largest. A partial restriction map indicates location between restriction sites. *lac z* is the region coding for *β* galactosidase; *lac i* is the region coding for the *lac* repressor. The starting point of transcription for the *lac* operon mRNA is designated as +1. CAP, P and O indicate the binding sites for the CAP–cyclic AMP complex, RNA polymerase and *lac* repressor respectively. Each fragment was cloned into *Eco*RI sites[21,22]. Purified fragments thus had four-base *Eco*RI ends. The influence of these single-stranded ends on theoretical melting curves was negligible.

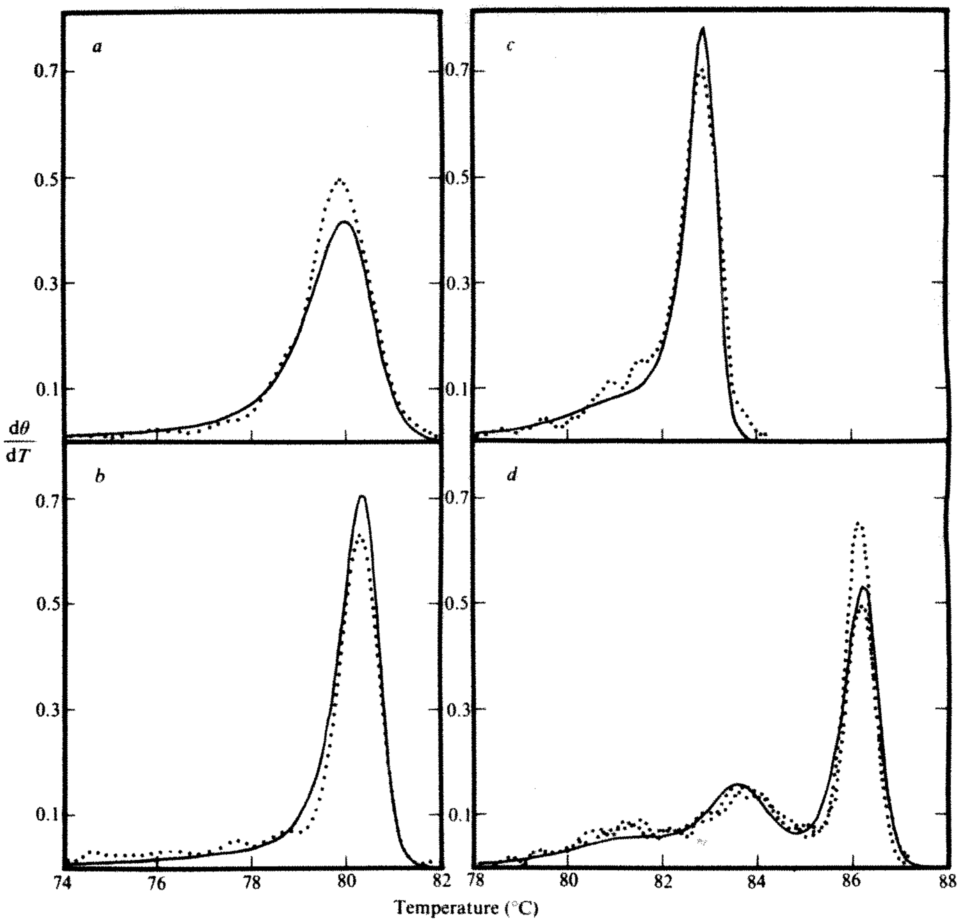


Fig. 2 Experimental differential melting curves of *lac* operon fragments (dotted lines) and theoretical curves (solid lines) are plotted. *a* Shows the 95-base pair DNA, *b*, the 144-base pair DNA, *c* the 203-base pair DNA and *d* the 301-base pair DNA. The solvent was 0.1 M NaCl, 1 mM potassium phosphate and 10 μM EDTA. Experimental data for the 95-, 203- and 301-base pair DNAs were from Hardies *et al.*[10]. The 144-base pair fragment was isolated in this laboratory using previously described methods[21,22]. Melting curves of the 144-base pair DNA were obtained in the same solvent as above. A double-beam spectrophotometer was used. A platinum resistance thermometer measured the sample temperature. The heating rate was 6 °C h^{-1} and absorbance readings at 260 nm were made at approximately 0.05 °C intervals. Data were collected on paper tape and then processed with a time-sharing computer facility. Detailed procedures will be described elsewhere (D.K.H. and R.M.W., unpublished). The data analysis is similar to that used by Vizard *et al.*[2]. The T_m values of the 95-, 203- and 301-base pair DNAs were reported in ref. 10. Reproducibility of these T_ms was within 1 °C. Duplicate experiments of the 144-base pair fragments gave a T_m of 80.9 ± 0.2 °C. *d* Shows the internal reproducibility of two melting curves of the 301-base pair fragment. The major peaks of the two dotted experimental curves have been aligned for comparison at the average T_m. Experimental error for the other three fragments was no greater than that shown in *d*. Calculated curves were shifted slightly in temperature (0.2–0.6 °C) to match the average T_ms of the major peaks. The small oscillatory peaks in the early part of the experimental transitions are probably due to the cubic fit smoothing of the raw absorbance data. We do not regard them as real intensity peaks. Melting curves of the 95- and 301-base pair fragments were also obtained at 282 nm. The 95-base pair curve was within the experimental error of melting transitions measured at 260 nm. The only significant difference between the 301-base pair curve at 282 nm and *d* was 20% decrease in the amplitude of the minor peak. Theoretical parameters used were $T_{AT} = 64.3$ °C, $T_{GC} = 107.2$ °C, $\sigma = 4.5 \times 10^{-5}$, $H_{AT} = -8.5$ kcal mol^{-1} and $H_{GC} = -9.4$ kcal mol^{-1}. The loop entropy term was that used by Wartell[23]. As expected, varying the loop entropy constant ($1.0 < k < 1.9$) did not alter the results. The dissociation parameter, β, is given in the text and Fig. 3 legend.

good fit to the melting curve of the 301-base pair DNA and β values for the other DNA fragments were determined from equation (2). This procedure provided very good agreement with the experimental T_m values and overall transition shape. However, the theoretical curves consistently predicted transition widths slightly broader than those observed experimentally. This discrepancy was removed by altering the simplified model to include dependence of the external partition functions on θ_{int}. The moments of inertia for various duplex configurations clearly depend on the amounts of base-paired and single-stranded regions. Figure 3 illustrates DNA configurations at different stages of melting. We assume that within the helix–coil transition region, DNA melting resembles a rod dissociating to two coiled spheres for the initial stages of melting and a sphere dissociating to two coiled spheres at the upper end of the transition. Thus α varies as $\alpha = a + b(1 - \theta_{\text{int}})$, where the constants a and b are chosen such that α decreases from -3 to -4 through the transition. This change removed the transition-width discrepancy.

Theoretical calculations can determine θ_i, the probability that the ith base pair of the sequence is intact at a given temperature[14,15]. Melting profiles—plots of $(1 - \theta_i)$ against i as a function of temperature—indicate the temperatures at which specific regions within the DNA sequence unwind. Note that θ_i depends only on a DNA's internal degrees of freedom, that is,

$$\left(\sum_{i=1}^{N} \theta_i\right) \Big/ N = \theta_{\text{int}}.$$

Figure 4 shows melting profiles of the 301-base pair fragment between 80 and 90 °C. Similar profiles have been obtained for the 95-, 144-, and 203-base pair fragments (not shown). The profiles indicate that denaturation initiates from the ends with regions of 60–80 base pairs melting over a 1–2 °C interval. These regions can sometimes be associated with peaks on experimental

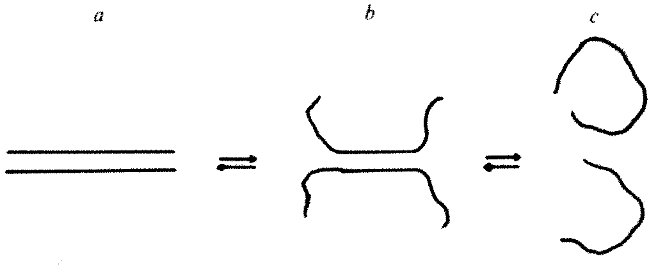


Fig. 3 Simple model for short DNA helix–coil dissociation. Three configurations are labelled *A* (helix rod), *B* (helix rod–sphere) and *C* (two coiled spheres). β, the ratio of partition functions for the external degrees of freedom, depends on fragment length N, and the fraction of bonded base pairs among duplexed strands, θ_{int}. $\beta = KN^{\alpha}$, where $\alpha = a + b(1 - \theta_{\text{int}})$. Constants K, a and b were 5×10^3, -2.8 and -3.2 respectively. a and b were chosen such that when combined with θ_{int}, α varied from -3 to -4 through the transition. In the early stages of the transition, $\theta_{\text{int}} \approx 0.9$, $\alpha \approx -3$ and dissociation is best described as $A \rightarrow C$. Base-pair unwinding with increasing temperature changes the configuration of the partially duplexed DNAs. Duplex dissociation resembles $B \rightarrow C$ as the temperature increases. For the last portion of the transition, $\theta_{\text{int}} \approx 0.6$ and $\alpha \approx -4$.

curves. Figure 4 shows the right-most 52 base pairs melting by 82 °C. This region corresponds to the slight increase from 80–82 °C in Fig. 2*d*. The small peak centred about 83.6 °C in Fig. 2*d* is associated with the melting of an 88-base pair block adjacent to the previously unwound region. These results are similar to those deduced by Hardies *et al.*[10] from analysis of the experimental curve and base-composition distribution of the 301-base pair fragment. One must be careful not to overinterpret melting profiles such as those in Fig. 4. As θ_i values do not consider strand dissociation, a meaningful comparison of melting profiles with differential melting curves is only possible when most strands are in the duplex state. Figure 4 shows an 80-base pair region largely intact at 88 °C, where the net fraction of intact base pairs, θ_{net}, is close to zero (Fig. 2*d*). At this temperature $\theta_{\text{int}} \approx 0.3$ but $\theta_{\text{ext}} \approx 10^{-5}$. Strand dissociation has reduced duplex DNAs to an insignificant fraction of total strands. Figure 4 would be valid throughout the transition if one could cross-link the complementary strands to prevent strand dissociation. Melting profiles of the 95-, 144- and 203-base pair fragments also show that melting initiates from the ends.

The agreement between theory and experiment indicates that the theoretical model is accurate for the conditions examined. Further studies on other short DNA restriction fragments also show good theory–experiment agreement (W. Hillen, T. C. Goodman, R. D. Wells, A.S.B. and R.M.W., unpublished results). It remains uncertain why exact calculations do not agree well with experimental curves of longer DNAs. The theory provides a method for examining the influence of base pair sequence on DNA melting of short fragments. Additional calculations show that single base pair changes in certain locations in the 203-base pair DNA can alter the overall melting temperature by 0.6 °C. This type of information may be relevant to understanding the interaction of RNA polymerase with promoter sites. Several results indicate that the thermal stability of promoters and their surrounding regions influence RNA polymerase–promoter interaction[18–20].

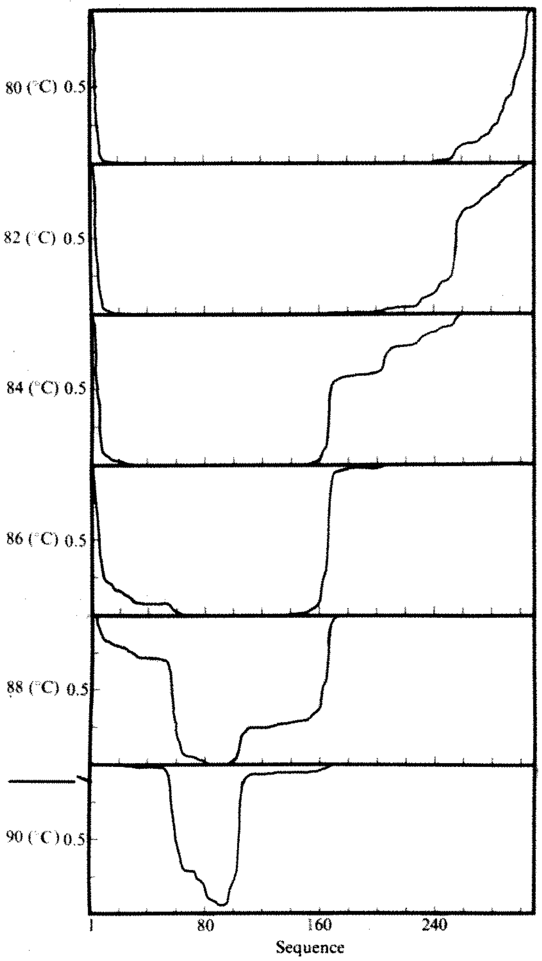


Fig. 4 Probability profile for the 301-base pair fragment. Probability of a base pair broken $(1-\theta_i)$ is plotted on the *y* ordinate against base pair (*i*) of the 301-base pair DNA sequence as a function of temperature.

We thank S. C. Hardies, W. Hillen, T. C. Goodman and R. D. Wells for providing us with their manuscript before publication, and for helpful discussions and supplementary information. This work was supported by the NSF grant PCM 76-04565, Biomedical Research Support grant 5-S07-RR07024, and a grant of computing time from the Office of Computing Services at the Georgia Institute of Technology. R.M.W. holds a Research Career Development Award of the NIH.

Received 16 April; accepted 13 October 1980.

1. Lyubchenko, Y. L., Vologodskii, A. V. & Frank-Kamenetskii, M. D. *Nature* **271**, 28–31 (1978).
2. Vizard, D. L., White, R. A. & Ansevin, A. T. *Nature* **275**, 250–251 (1978).
3. Wada, A., Tachibana, H., Ueno, A., Husimi, V. & Machida, Y. *Nature* **269**, 352–353 (1977).
4. Tachibana, H., Wada, A., Gotoh, O. & Takanami, M. *Biochim. biophys. Acta* **517**, 319–328 (1978).
5. Gabbarro-Arpa, J., Tougard, P. & Reiss, C. *Nature* **280**, 515–517 (1979).
6. Tong, B. T. & Battersby, S. J. *Biopolymers* **18**, 1917–1936 (1979).
7. Ueno, S., Tachibana, H., Husimi, Y. & Wada, A. *J. Biochem.* **84**, 917–924 (1978).
8. Azbel, M. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 105 (1979).
9. Azbel, M. Y. *Biopolymers* **19**, 61–109 (1980).
10. Hardies, S. C., Hillen, W., Goodman, T. C. & Wells, R. D. *J. biol. Chem.* **254**, 10128–10134 (1979).
11. Dickson, R. C., Abelson, J., Barnes, W. M. & Reznikoff, W. S. *Science* **187**, 27–35 (1975).
12. Lyubchenko, Y. L., Frank-Kamenetskii, M. D., Vologodskii, A. V., Luzurkin, Y. & Gause, G. G. Jr *Biopolymers* **15**, 1019–1036 (1976).
13. Poland, D. & Scheraga, H. A. *Theory of Helix–Coil Transitions in Biopolymers* (Academic, New York, 1970).
14. Wartell, R. M. & Montroll, E. W. *Adv. chem. Phys.* **22**, 129–203 (1972).
15. Poland, D. *Biopolymers* **13**, 1859–1871 (1974).
16. Applequist, J. & Damle V. *J. chem. Phys.* **39**, 10 (1963).
17. Landau, L. D. & Lifshitz, E. M. *Statistical Physics* (Addison-Wesley, Reading, Massachusetts, 1958).
18. Botchan, P. *J. molec. Biol.* **105**, 161 (1976).
19. Jones, B. B., Chan, H., Rothstein, S., Wells, R. D. & Reznikoff, W. S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4914 (1977).
20. Vollenweider, H. J., Fiandt, M. & Szybalski, W. *Science* **205**, 508 (1979).
21. Hardies, S. C. & Wells, R. D. *Gene* **7**, 1–14 (1979).
22. Wartell, R. M. & Reznikoff, W. S. *Gene* (in the press).
23. Wartell, R. M. *Nucleic Acids Res.* **4**, 2719 (1977).

Detection of the effects of asymmetric assortative mating

L. J. Eaves & A. C. Heath

Department of Experimental Psychology, South Parks Road,
Oxford OX1 3UD, UK

Assortative mating is clearly important in determining the genetic structure of a population but studies of assortative mating in man have been restricted mainly to the examination of correlations between spouses for single variables. It has been assumed that assortation is based directly on the measured phenotype or on components of the phenotype whose determination is similar in both sexes. Recently, such assumptions have been questioned (see, for example, refs 1–3) because males and females may regard different aspects of the phenotype as important for mate selection. The term 'asymmetric assortative mating' has been used to describe such situations. It has also been suggested that the study of the spouses and offspring of twins[1,2] may permit resolution of such components of the mating system. *Ad hoc* models have been proposed (for example, ref. 1) to specify the contribution of asymmetric assortment to the correlations between relatives, but these have not been internally consistent or sufficiently parsimonious. We show here how a simple model based on biological considerations offers a foundation for a consistent and economical theory of asymmetric mate selection which may be used for the analysis of patterns of familial similarity.

0028-0836/81/030205—02$01.00

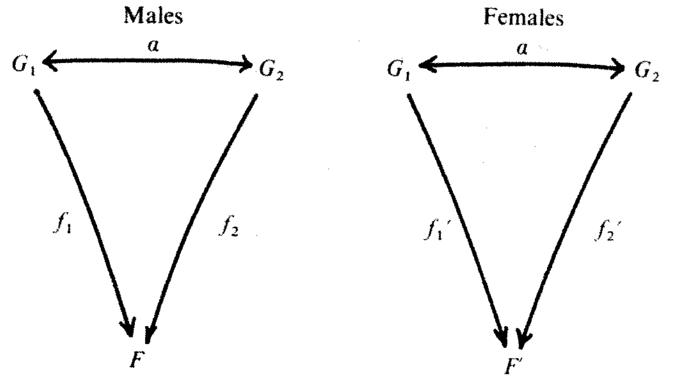


Fig. 1 The contributions of two genetic systems to variation in male and female fitness. See text for definition of symbols.

Table 1 The genetic correlation between MZ twins and their relatives in the presence of asymmetric assortative mating for fitness

Genetic correlation	Sex of twins	
	Male	Female
Twin	1	1
Twin–spouse	$\phi\phi'\mu$	$\phi\phi'\mu$
Spouse–spouse	$\phi'^2\mu^2$	$\phi^2\mu^2$
Sibling	$\frac{1}{2}(1+\phi\phi'\mu)$	$\frac{1}{2}(1+\phi\phi'\mu)$
Half-sibling	$\frac{1}{4}(1+2\phi\phi'\mu+\phi'^2\mu^2)$	$\frac{1}{4}(1+2\phi\phi'\mu+\phi^2\mu^2)$
Parent–offspring	$\frac{1}{2}(1+\phi\phi'\mu)$	$\frac{1}{2}(1+\phi\phi'\mu)$
Spouse–nephew/niece	$\frac{1}{2}(\phi'^2\mu^2+\phi\phi'\mu)$	$\frac{1}{2}(\phi^2\mu^2+\phi\phi'\mu)$

See text for definition of parameters.

The model assumes that assortative mating is based on a composite aspect of the phenotype, 'fitness', and is necessarily symmetric at this level. Asymmetry in mate selection arises because the sexes differ in the contribution of particular genetic effects to this composite phenotype, thus the determination of fitness is partly sex-limited. The principles are illustrated by a simple special case in which there is no environmental variation in fitness. The model can be extended to incorporate such additional subtleties as maternal effects and cultural transmission without altering its essential features.

For simplicity we postulate that fitness in males, F, and in females, F', is determined by two genetic systems, G_1 and G_2, which are correlated within individuals simply through the genetic consequences of the mating system. The contributions of G_1 and G_2 to variation in F and F' are represented by the standardized regression (path) coefficients f_1 and f_2 for males and f_1' and f_2' in females. The correlation between G_1 and G_2 is denoted by α. The possibility of asymmetry is represented by the fact that the paths in males and females, f_1 and f_1', f_2 and f_2' are not necessarily the same. Figure 1 represents the model in diagrammatic form.

It is assumed that assortation is based primarily on the phenotypic correlation (which is symmetric by definition) between F and F'. This is the marital correlation, μ, for fitness. The critical determinant of the distribution of genetic differences in the next generation is the genetic correlation between spouses. We consider that between the deviations in G_1 for males and females. The algebra is equivalent for any other component of F and F', whose effects may be measured directly or indirectly. Given α, μ, f_1, f_2, f_1' and f_2', the correlation between the G_1s of spouses is $\mu\phi\phi'$ where $\phi = f_1 + f_2\alpha$, and $\phi' = f_1' + f_2'\alpha$. As both ϕ and ϕ' are involved in this correlation, they will not be resolved by examining the phenotypic correlation between spouses for a single variable. Such resolution is possible, however, if we consider the spouses of relatives. Following Nance[1] and Eaves[2], we consider the spouses of identical (MZ) male and female twins. To explore the consequences of asymmetry for the next generation we also consider the genetic correlations between siblings and the biological half-siblings obtained by mating MZ twins to separate spouses. We consider only the effects of G_1, as shown in the model in Fig. 2. The coefficients of $\frac{1}{2}$ on the paths between parental and offspring genotypic deviations reflect an underlying assumption of genetic additivity. The diagram assumes that the twins are female. The figure for male twins would be similar except that ϕ would be replaced by ϕ' and vice versa.

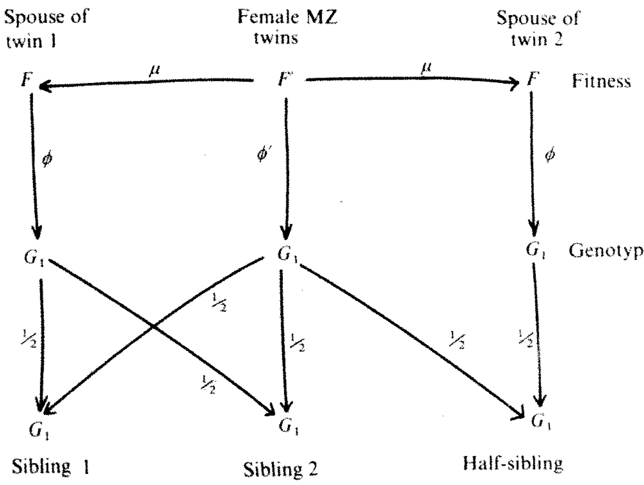


Fig. 2 The contribution of assortative mating to the genetic similarity of the relatives of female identical twins.

The expected genetic correlations between relatives are given in Table 1. These can be obtained by applying the rules of path analysis to the model in Fig. 2. The sex of the twins does not affect the sibling and parent–offspring correlations differently. The maternal and paternal half-sibling genetic correlations differ by $\mu^2(\phi - \phi')$ and that between the spouses of female and male twins by $\mu^2(\phi^2 - \phi'^2)$. As Nance[1] suggested, the effects of asymmetry may indeed simulate those of classical maternal effects in half-sibship data by making maternal half-siblings more alike than paternal half-siblings if $\phi > \phi'$. It is the comparison of the spouses of twins which may have a critical role in the resolution of direct maternal effects from those created or simulated by assymmetric mate selection. Environmental components in the determination of F and F' change the definition of ϕ and ϕ' without substantially altering these conclusions. Clearly, resolution of genetic and environmental determinants of mating preference requires data on the spouses of relatives in addition to MZ twins[2], but the model we propose may form the basis for the discussion of more complex systems of mating than those which are frequently considered by students of human differences.

A. C. H. holds a MRC studentship. The work is part of a programme supported by the MRC and of a project supported by the SRC.

Received 9 May; accepted 3 November 1980.

1. Nance, W. E., In Sing, C. F. & Skolnick, M. in *Genetic Analysis of Common Diseases: Applications to Predictive Factors in Coronary Disease*, 453–464 (Liss, New York, 1979).
2. Eaves, L. J. *Heredity* **43**, 399–409 (1979).
3. Rose, R. J. *et al. Nature* **283**, 375–377 (1980).

BOOK REVIEWS

Anti-realist philosophy of science

Mary Hesse

The Scientific Image. By Bas C. van Fraassen. Pp.235. (Clarendon/Oxford University Press: 1980.) Hbk £15, $45; pbk £6.50, $18.95.

PHILOSOPHERS of science have recently been much exercised about the question "Do scientific theories describe the real world?". This is a question that arises from two sorts of problem. First, it has become a question within quantum physics, and has led to disputed interpretations such as those between the so-called Copenhagen school and the "hidden variables" theorists. Second, it has become a matter of particular concern to philosophers since the radical analyses of Kuhn and Feyerabend in the 1960s, in which it was suggested that the contents of particular scientific theories are culturally relative pictures of the world, not attempts at true objective explanation. The villagers of Salem believed in witches, modern scientists believe in atoms; both are unobservable, so why should the latter be considered to be nearer the truth? Even to raise such a question constituted an attack upon the rationality and progressiveness of science, and sent philosophers scurrying back to the defence of realism.

Van Fraassen's book marks a significant new stage in this debate. In it he defends an anti-realist position, but without falling into the extreme absurdities of relativism. He draws the lines of the debate in the following way. What he calls "minimal realism" holds that "science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true" (p. 8). What he calls "constructive empiricism", on the other hand, holds that "science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate" (p. 12). His arguments against the former view and for the latter are detailed and cogent, and range over the problematic relation of theory and observation, the nature of scientific explanation and scientific methodology, and the proper interpretation of "probability" in statistical science. I will pick out a few of the salient and more controversial points of the argument.

First, as van Fraassen construes the debate, it is about the aims of science, not about the nature of theories as such. He does not claim that theories are "nothing but" computing devices for processing and emitting predictions, nor that they are metaphorical fictions which speak only and misleadingly about observations. For van Fraassen, theories consist of meaningful propositions having truth value. They are given meaning in the same way that any descriptive language is given meaning (whatever that is). Van Fraassen does not consider it the job of philosophy of science to debate general theories of meaning, and here he reveals a belief that science and the philosophy of science do not fundamentally upset the *status quo* with regard to received theories of language, meaning and truth. Thus he does not follow those radical philosophers of science who consider that scientific concepts may so modify our commonsense picture of the world as to require changes also in the way we conceive natural language. van Fraassen does not believe that philosophy of science has implications for "deep" philosophical problems, or that philosophy of science is properly about "language" at all — it is about the world, just as science itself is about the world. Such conservative assumptions are controversial, but van Fraassen is certainly correct in seeing that if they were to be questioned, we would be led into issues far beyond the range of this book.

Theories, then, have truth value, but it is not the aim of science to discover true theories. Though it is not argued quite explicitly in the book, this rejection of theoretical "explanation" as the chief aim of science seems to be based on the belief that we can never know that we have arrived at exactly true theories. Most theories have turned or will turn out to be false as more evidence comes in, and even where they have not yet been refuted, there are always conceivable incompatible alternatives which are also empirically adequate. Thus the aim of discovery of true theories is one that we cannot ever know that we have attained, and it should be replaced by the realizable and often realized aim of empirical adequacy. To the currently fashionable argument that theories must be true of the real world because they are so successful in application, van Fraassen simply replies that they are successful because they incorporate accounts of real empirical regularities; it is these that make them successful, and only those theories that incorporate them survive. But we have no warrant to claim truth for what goes beyond empirical regularities, even though all interesting theories do undoubtedly go beyond them.

Other common arguments for and against realism have started from the so-called "theory-ladenness" of observation. On behalf of realism it is argued that since our theoretical view of the world reaches right down into the way we describe observations ("tracks of *electrons*", "*glacial* moraines") and into the expectations we form regarding possible future observations, therefore we are bound to presuppose that at least some of our background theories are true. On the other hand, the anti-realist replies that this all-pervasive character of theory simply shows that our theoretical world-views are cultural products, which can be seen in historical and social perspective to change from culture to culture, and are therefore culturally relative and not realist. Van Fraassen makes short work of all this argumentation. Our language, both theoretical and observational, is of course the language of *our* culture group. But this does not make observational *information* either theory- or culture-relative — only its *expression* as embodied in the current scientific language. Theories are not linguistic entities; they should rather be seen as families of models of possible worlds, whose empirical sub-structures represent phenomena.

Talk of models commits van Fraassen to a metatheoretical view of science in the sense that he presupposes that theories are ideally axiomatized, and that their empirical content is exhausted by the empirical interpretations of entailments of the axioms. This view results in what I would regard as a major limitation of the book, namely van Fraassen's treatment of scientific methodology. He sets out to show that accepted methodological principles such as simplicity, explanatory power, the unity of science and so on are explicable in terms of an anti-realist view. In some cases he succeeds, but in others he has to be content to regard the maxims of theory-choice as "merely pragmatic", that is to say not relevant to questions of truth or the empirical adequacy of predictions. In terms of his position he cannot answer questions such as: why should predictions from simpler theories be generally more reliable?; why should analogical arguments from theory to theory work as well as they do? I believe van Fraassen's comparatively cavalier dismissal of such problems of epistemology, induction and confirmation can be traced to his exclusively deductive (and sometimes statistical) construal of the formal character of theories. But theories are not treated in scientific practice as axiomatized entities; they have much fuzzier edges than that, and reliably support inferences that are not just deductive or formally statistical.

Such limitations apart, however, this book for the first time makes a clear statement of what is involved in scientific realism, and it will become the classic current argument against that philosophical position. It successfully debunks recent exaggerated attention to the concept of "explanation", and replaces it with a healthy concern with truth and empirical adequacy that does not give any hostages to relativism. Tactically, its relative conservatism about the general philosophical implications of philosophy of science will make it all the more influential among realist philosophers. Scientists, however, may well be ready for more radical investigation of the effects of changing scientific theories on our commonsense presuppositions, and particularly on our views of logic, natural language and scientific inference.

Mary Hesse is Professor of Philosophy of Science, University of Cambridge.

Archaeological jargon explained

A. F. Harding

A Dictionary of Terms and Techniques in Archaeology. By Sara Champion. Pp.144. (Phaidon/Facts on File: 1980.) Hbk £6.95, $15.95; pbk £4.95.

THE Introduction to this little volume tells us that the book "is not a general dictionary of archaeology: names of cultures and artefacts have been deliberately omitted". Instead, it "has been compiled with the aim of introducing to non-professional archaeologists the terms and techniques used in modern archaeology". The author disclaims personal expertise in most of the techniques described, and has generally provided brief (typically 30 to 300 word) entries with a reference to a standard work on the subject.

Despite its limited brief, professionals and students, as well as interested amateurs, will find much of practical use in this book, especially for checking up on those jargon terms one so frequently hears but secretly wonders about. The *Dictionary* contains only about 330 entries, many of them illustrated, so the scope is necessarily selective: most space is devoted to modern archaeological theory, scientific techniques in archaeology, archaeological techniques *per se,* terms from related disciplines and a certain number of technological terms — especially those relating to potting and metalworking. Stone and stone-working, to take a single example, are barely touched upon.

Each reader will undoubtedly find his own lacunae, and it is a pity that the author was not able to roam more freely. Yet the book is well-produced and within its self-imposed limits can be recommended.

A. F. Harding is a Lecturer in Archaeology at the University of Durham.

The biology of history

Hugh Thomas

The Human Condition. By W. H. McNeill. Pp.75. (Princeton University Press: 1980.) £4.75, $11.25.

UNIVERSAL histories are becoming alarmingly fashionable. This is understandable for three reasons. First, although we are a long way still from anything like a world state we are quite close to an interdependent world society. The slightest hiccup in southern Asia affects northern Europe. Oil, especially, is the great internationalizing force, causing all responsible people to remember statistics of production in countries whose very names were unknown to them ten years ago. Second, the concentration of such enormous power in the hands of two superpowers is, also, a development which has never occurred before in human history: until 1945, power was always far more widely dispersed. The third explanation is that the apparent long-term shortages in raw materials have persuaded human beings to reflect seriously over the springs of their predicament in a new way. There is no doubt but that this is a benefit. National revival is unilaterally possible (such as occurred in France in 1958) but it is unlikely to occur save in the climate of international prosperity.

W. H. McNeill is no newcomer to the grand approach. His *Rise of the West* (Chicago University Press, 1963) is rightly regarded as one of the best, probably the best, of the conventional outline general histories. His *Plagues and Peoples* (Blackwell, 1977) was an extremely interesting introduction to the role of disease in history. *The Human Condition* is more modest, being merely two (long) lectures given at Clark University. Or at first sight modest, because, after all, there is a sense in which any book with this title which seeks adequately to cope with such a subject-matter in 75 pages is magnificently ambitious.

The themes of Professor McNeill's essay (for such it is) derive from his earlier study of disease. The main concept is what he calls "parasitism", divided in its challenge to human beings into "macroparasites" and "microparasites". Microparasitism refers to the "metabolic activities of minute organisms" which compete with human beings for food: by getting there first, microparasites can forestall human efforts to capture energy from food. Macroparasitism is any "exploitative relations among groups and classes of human beings", such as when one man or group of men — armed raiders, tax collectors, rent collectors — seize goods or compel services from other groups.

This rather forbidding frame gives Professor McNeill a justification for what turns out to be an extremely lively and well-written short world history, with the use of little jargon other than that I have already quoted. The frame enables Professor McNeill to make some interesting and novel points: for example, he attributes the advance of our remotest ancestors to the "apex of the food chain" to the acquisition of language and to the superior co-ordination of human behaviour that language allowed. The next landmark was the use of clothes to "maintain a tropical microenvironment" next to the almost hairless skin we all inherit. A bearskin on the back enabled us to live in a cold climate. Agriculture, though a definite benefit, Professor McNeill somewhat grudgingly confirms, brought microparasites as well as harvests. Irrigation brought mosquitoes. The "first recorded conqueror" Sargon of Akkad (*c.*2250 BC) was the first macroparasite on an imperial scale. Civil servants were the consequence — a problem with which we are still wrestling. In the sixteenth century, a fundamental change occurred, worthy of greater attention than Professor McNeill gives it — businessmen succeeded in establishing real autonomy from their political superiors, with even the mightiest European rulers depending on traders and bankers to organize an efficient army. They had to solicit loans from private persons who were usually domiciled outside the ruler's own jurisdiction and so were safe from confiscatory taxation. This really began the modern age — despite Professor McNeill's encouragement to us to look at the Sung dynasty in China AD 960 – 1279 in this light. (Incidentally, there is a certain discrepancy between Professor McNeill's comments on mediaeval China — p.48 — and his comparison with Renaissance Europe — p.62; he seems really to stand by the latter interpretation where he says that private businessmen in China never won autonomy from their superiors nor accumulated capital on the scale that European entrepreneurs did.)

The last pages of Professor McNeill's book are provocative and interesting and end on a sensible note: "better hindsight deepens insight". Foresight also will be assisted by all sensible enough to get hold of this excellent little book.

Hugh Thomas was formerly Professor of History at the University of Reading.

Alternative energy: the biological option

C.R.W. Spedding

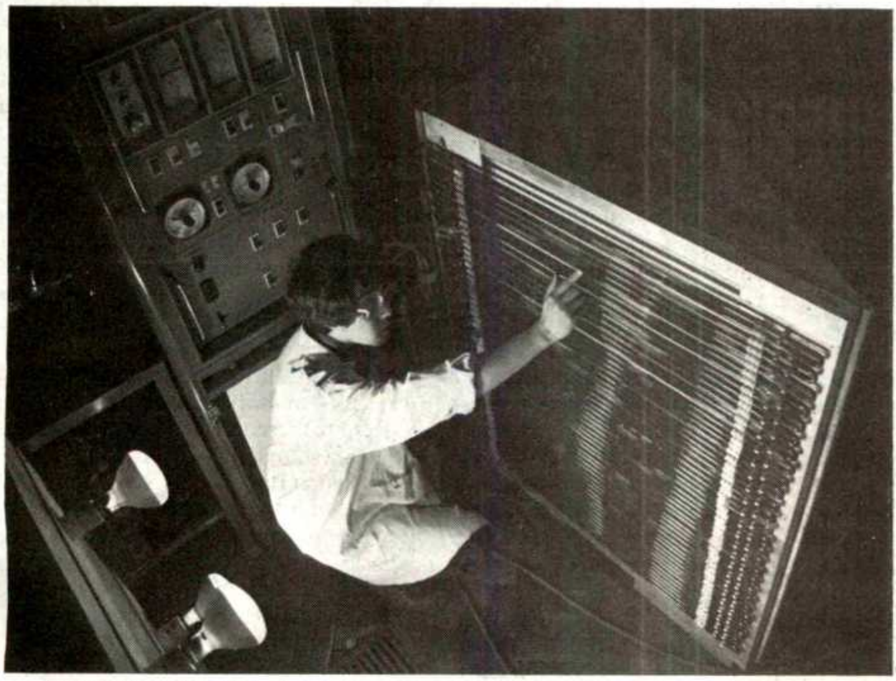
Biochemical and Photosynthetic Aspects of Energy Production. Edited by Anthony San Pietro. Pp.231. (Academic: 1980.) £15.80, $28.

AS THE editor of this book points out in his preface, photosynthesis is the only method of solar energy conversion currently practised on a large scale. Increasing dependence on fossil fuels is now an extremely vulnerable condition, and this volume explores the potential of biological processes that may serve in the future to provide alternative energy resources.

Eleven contributors discuss, in considerable detail, various aspects of this enormous subject. D.O. Hall provides an introductory overview and reviews the development of photobiological conversion systems and their long-term implications from the point of view of both energy and food. He stresses the need to consider all our energy options and not repeat the mistake of putting all our eggs in one basket.

Four chapters deal with different forms of energy from biomass production systems based on seaweed aquaculture, algal–bacterial systems, anaerobic fermentation of residues and biomass, and energy crops. All of these need further exploration; some are not yet established as technically feasible and all depend for their relevance on relative costs. This presents a major problem: how to estimate costs and prices for the future for technologies that are currently at a stage equivalent to the first motor car or aeroplane. The value and role of energy analysis is considered, but this cannot be a complete answer without regard to the fact that one joule of energy is not necessarily of the same value as another.

Several contributors emphasize the importance of non-energy products in our assessments. In some cases, such as glycerol production from the alga *Dunaliella* spp., the output is predominantly of one substance, but the economics of most processes will depend upon a use for all of the fractions produced.



Still small — photobiological conversion systems. An experimental tubular loop reactor which converts artificial solar energy into biomass.

Courtesy British Petroleum/Professor S.J. Pirt, Queen Elizabeth College

Other issues considered are the atmospheric consequences of various developments and the biological fixation of nitrogen. The latter is of key importance to greater and energetically more efficient production of biomass.

The book is well produced, has a useful index and comprehensive lists of references, and represents a valuable summary of the state of an important subject.

C.R.W. Spedding is Professor of Agricultural Systems and Head of the Department of Agriculture and Horticulture in the University of Reading.

Polyamines reviewed

Anthony E. Pegg
Hannu Pösö

Polyamines in Biomedical Research. Edited by Joseph M. Gaugas. Pp.474. (Wiley-Interscience: 1980.) £30, $90.

KNOWLEDGE of the importance of polyamines in cellular physiology has been greatly extended in recent years by studies of the control of polyamine biosynthesis, the development of drugs interfering with polyamine accumulation and improved technology for measurement of these compounds and derivatives. Many reviews on these subjects have appeared during the past few years and this book is the latest in the series. It contains a number of outstanding chapters and, though dated, will be of considerable interest to those working in the field.

The book is a collection of 27 review chapters by different authors. About one-third of these are concerned with the control of polyamine synthesis and the role of polyamines in cell growth. There are naturally substantial overlaps in which different authors describe results from the same publications but all of these articles are by experts in the field and are of interest. Chapters by Stevens and Stevens and by Mamont, Bey and Koch-Weser give comprehensive and well-written coverage of inhibitors of decarboxylases involved in polyamine biosynthesis. This provides excellent reference material, as does the contribution on regulation of ornithine

Erratum: see *Nature* **288**, 630. The title of the book written by Sandra Raphael and Wilfrid Blunt is *The Illustrated Herbal* (Frances Lincoln/Thames & Hudson, £12.50, $25).

decarboxylase by McCann. The possible importance of polyamines and ornithine decarboxylase in skin tumour promotion and as markers for compounds with anti-promoting actions are well described by O'Brien and by Verma and Boutwell who were pioneers in the field. There is also a first-rate and — at the time of writing — comprehensive evaluation of methods for polyamine measurement by Seiler who is responsible for much of the methodology now in routine use. Isolation and analysis of novel polyamines found in thermophiles are covered in a chapter by Zappia and colleagues. The remainder of the book contains a heterogeneous collection of contributions on carcinogenesis, polyamine degradation, their putative roles in macrophage function, relationship to chalones, suppression of lymphocyte function, excretion and conjugation, and value as tumour markers. These articles are often highly speculative and, while thought provoking, are much less rigorous.

Several important contributors to polyamine research are missing from this volume (for example it is unfortunate that there is no chapter from H. and C.W. Tabor on their polyamine-deficient mutants), and the two-year delay between writing and publication has meant there are a number of significant omissions in content. However, as a source book for scholars of a wide range of interests, it should prove valuable.

Anthony E. Pegg is Professor in the Department of Physiology at The Milton S. Hershey Medical Center, Hershey, Pennsylvania. Hannu Pösö is Docent in the Department of Biochemistry at the University of Helsinki.

Air of learning

Peter S. Liss

Air in Danger: Ecological Perspectives of the Atmosphere. By G. Breuer. Pp.189. (Cambridge University Press: 1980.) Hbk £10, $24.95; pbk £3.50, $7.95.

IMAGINE you are a scientist or scientific journalist invited to attend a research meeting on a topic which is not your speciality. You enjoy the meeting and become sufficiently interested in the topic to follow it up by reading, and visiting and writing to experts in the field. Then you decide to write up not only what you have learnt but also the various stages in the learning process, and persuade major publishers to issue the result in both German and English.

Although no biographical details of the author of *Air in Danger* are given (and, for once, I felt their absence to be a real loss), this would appear to be the way in which this book came into being. Georg Breuer attended, as an observer, the Dahlem Conference, "Global Chemical Cycles and Their Alterations by Man", held in Berlin in November, 1976. This spurred him to find out more about a major topic of the meeting, the way in which the major atmospheric gases — nitrogen, oxygen and carbon dioxide — are cycled in the environment. The chapters of the book chronicle the author's education as to how normal biological processes transfer these elements between the various environmental reservoirs and also how humans are influencing this cycling. Quite properly, considerable attention is given to the photosynthesis/respiration–decomposition reactions since these truly dominate the environmental chemistry of the gases considered, as well as that of many other substances.

Inevitably, this way of setting about writing a book has its disadvantages. Someone who is self-taught spends considerable time exploring what turn out to be blind alleys (for example, several pages in the first chapter deal with what is well known to be the "non-problem" of whether fossil fuel combustion can significantly deplete atmospheric oxygen) and also tends to over-detailed explanations.

On the plus side the approach adopted has the freshness inherent in knowledge newly acquired. What is not clear is what type of reader will want to share in the author's learning experience, since each of us learn in an individualistic and idiosyncratic manner. The book is certainly timely in that it presents in considerable detail much of the background behind the current debate over Lovelock's Gaia hypothesis, recently expounded in his book *Gaia: A New Look at Life on Earth* (Oxford University Press, 1979 — see *Nature* **282**, 154 for review).

The translation (by Peter Fabian) from the German original reads very well, but the English edition should never have been titled as it is. The main title gives the impression of a gloom and doom, pollution-orientated book which it certainly isn't. In fact, the author is to be commended for treading a very sensible line between natural and anthropogenic effects and the book's main title does his care a considerable injustice.

Peter S. Liss is Reader in the School of Environmental Sciences, University of East Anglia, and Scientific Advisor in the Environmental Studies Section at CEGB Headquarters.

Stimulating energy transfer

A.J. McCaffery

Introduction to Molecular Energy Transfer. By James T. Yardley. Pp.308. (Academic: 1980.) $34, £19.20.

THE microscopic processes of energy transfer — the molecular-level events in the gas, liquid and solid phases of matter — have been the subject of intense research interest for a number of years, and a book which attempts to draw together the wide-ranging strands of this field is most welcome. The topics covered in this volume reflect both the author's own research interests and the major thrusts of recent research, and thus the bulk of the book deals with vibrational energy transfer. Some of the modern experimental methods are reviewed in detail, as, more briefly, are the classical methods of ultrasonic dispersion and shock tube techniques. Gas phase vibration-to-translation and vibration-to-vibration transfer both receive detailed treatment and there is a short chapter on vibrational relaxation in liquids and inert materials. Included as a basis for the treatment of vibrational processes is a comprehensive theoretical development of the internal vibrational modes of molecules.

The remainder of the book is taken up with discussion of a diverse and interesting collection of topics. These range from radiationless electronic transfer in molecules, to the less known phenomena of collision-induced fine structure transitions in atoms and quenching of atomic excitation by curve crossing.

One interesting feature is a derivation of the Born approximation as early on as p. 8, and its subsequent usage runs like a thread through the wide range of fields covered in the later chapters, providing a useful physical insight into the individual transfer processes.

The difficulties of producing a completely up-to-date text are apparent, though references to work as late as 1978 are relatively frequent in some sections. Perhaps a more substantial criticism is that the book's purpose is not entirely clear. It is a mixture of pedagogy, in that it covers such basics as group theory, the hydrogen atom, LCAO-MO model and determinantal basis functions, and review. The coverage of rotational energy transfer is slightly disappointing by comparison to the treatment of its vibrational counterpart, as is the absence of discussion of scattering theoretic, semi-classical and classical trajectory studies of energy transfer.

Despite these small criticisms, Yardley's book is a stimulating and readable account of an important research topic with a number of interesting and novel features in its development. It should prove popular as a basic graduate text for North American universities and as a valuable source of methods and theory to research workers.

A.J. McCaffery is Reader in Molecular Sciences at Sussex University whose research interests are in the study of gas phase molecular dynamics using laser techniques.

nature

22 January 1981

Towards what shining city, which hill?

The American presidency is an impossible job which has an enlivening effect on those who do it. New encumbents embark on tasks widely held to be beyond the resources of one man with the zeal and enthusiasm of an army. Both Presidents Kennedy and Johnson, when new to office, offered the promise of a fresh start on old problems, domestic and international. Given the persistent importance of the United States in the affairs of most other communities, the promise of a change of direction echoes around the world. Some of the consequences are unsettling; changes of direction are disconcerting. Some are also enlivening, as for the new incumbent himself. It is too soon to know what will be the consequences of Mr Reagan's accession to the presidency, but a single inauguration speech is not a sufficient guide to what will follow in the months ahead. The safest assumption is that Mr Reagan, like all his predecessors, will be good at some things and less good, or even frankly bad, at others. There is no shame in that, given that the job is everywhere acknowledged to be impossible. After a spell in which the president seemed to be bad — but not for want of trying to be good — at most things, it is no wonder that Mr Reagan's coming has cheered even his political opponents. For a time, at least, people will be prepared to give President Reagan the benefit of the doubt. The vacillations and exaggerations of the election campaign will be forgotten. To that rule, there will be only one exception. In his acceptance speech in Detroit last summer, Mr Ronald Reagan promised that he would lead the American people to a "shining city on a hill". That phrase, redolent of Blake or Bunyan, is too arresting to forget.

President Reagan, no doubt less innocent than he seems, will not be starved of advice, solicited and otherwise. Politics being what it is, his presidency may well be made or broken in the next few months by almost accidental issues — the inflation rate, or the problems of Central America. In these connections, no new president would choose to start from where President Reagan must. One of the perils of the job is that he has no choice. Another is that the issues, although recently made clear, are by no means irrelevant to what may happen on the wider stage. If inflation continues as it has been in the past few years, American institutions will be weakened as surely as they have been undermined elsewhere, then the foundations of the shining city will be found not to exist. Already many of the new occupants of offices in the White House must be regretting that there had to be a campaign, and a promise of a tax-cut to go with it. If the problems of democracy in Latin America are ignored or, worse, supposed to be familiar problems susceptible to oldfashioned solutions, the roof could fall in. Even President Reagan's friends will be well aware of the seriousness of these dangers. Those who heard his acceptance speech may be forgiven for looking further ahead.

The Reagan position, so often defined in the past several months that it has become unclear, is ambiguous on two crucial issues — arms control and academics. In the parochial spirit in which trade magazines such as *Nature* grind their own axes on occasions such as the inauguration of a new president, special pleading may be excusable. On the face of things, for example, there is no reason why the new Adminstration should give an instant's thought to the special interest of the scientific community in the management of the arms race. Scientists, the argument might go, have (among other things) a special competence in military research and development. But only taxpayers have a right to say what should be done about relationships between the superpowers. President Reagan has in any case come a long way since the beginning of the election campaign, when he was bent on seeming a 1950s hawk. Now (the argument continues), he is for renegotiation of Salt II and other accommodations of that kind. The argument misses a point that will quickly become apparent, that the special but not exclusive interest of the scientific community in arms control is not merely academic. There is a sense in which J. Robert Oppenheimer was right in saying that "the physicists have known guilt" — hawks among them have persuaded doveish graduate students to work on unwelcome projects since time immemorial. The point has been reached at which both hawks and doves share with the general public the belief that it would be best if the consequences of the products of their joint endeavours could be contained within bounds, and that more or less any bounds would be better than none. This does not imply that all scientists share with the *Bulletin of the Atomic Scientists* the view that international tensions have worsened in the past few weeks — the *Bulletin* advanced its notorious minute hand one minute last week (and was mentioned on the British Broadcasting Corporation's morning news programme as a consequence) but did not convincingly explain why it had done so. Yet there is a great majority in the scientific community that would wish something sensible done about arms control, and which knows that such a course is not unattainable. The political question for the new president is whether he can see this as an opportunity.

Another issue for President Reagan, not too long to be left on the back burner, is that of the research community itself. Mr Carter (or, more probably, Dr Frank Press) has left the incumbent President with an awkward choice. Should an expansionary science budget be cut in the interests of financial prudence, or left as it is so as not to give offence? Neither of these questions should bother President Reagan in these early heady days. There is a simpler question he must answer. Since his early days as Governor of California, he has cut the figure of an adversary of the universities. At one stage, he forced Mr Clark Kerr out of his post as president of the University of California. Since then, he has made clear his disdain for academic preoccupations — decision is what matters. Such a position can be defended but is not defensible. The President needs urgently to build a bridge to the academics. To fail to do so will be to alienate some of those who might build a city, or find a hill.

Responsibility for trust in research

The public complaint, fashionable in the 1970s, that scientists could not be trusted when making statements about the hazards of pollution or of genetic manipulation, is now mercifully abating. Is the complaint now about to be replaced with the suspicion that scientists cannot trust each other? This will be some people's first reaction to the article by Harris *et al.* on page 228, and our Washington Correspondent's commentary on it (page 227). The reaction is understandable, even forgivable. For what has emerged is that three out of four laboratory cultures of cells originally described in 1977 as derived from the spleens of patients

0028-0836/81/040211-02$01.00

with Hodgkin's disease have nothing to do with human malignancy but are, instead, almost certainly derived from owl-monkey kidney tissue. This improbable conclusion has been established by means of as neat a piece of detective work as can be found in the recent literature. The fact remains that the four cell cultures have been the basis, in the past four years, for a spattering of spurious data in the scientific literature. Some of these unwarranted conclusions have been used by others in unrelated work towards the understanding of Hodgkin's disease. Elsewhere, laboratories have tried (usually without success) to establish their own analogous cell cultures from patients with the disease. But Dr John Long's apparent success with his supposedly authentic cultures appears to have led to the award of at least one substantial research grant (for $500,000 over three years), afterwards rescinded. Especially because the principal investigator had been accused of falsifying data in a different context, many will now be tempted to suppose that four years of work with Hodgkin's disease has been a kind of hoax.

To jump to that conclusion would be wrong. Those who work with cells in culture are continually aware of the risks of contamination. Especially in laboratories specializing in this work, one established cell culture may be contaminated with cells from another by a variety of means — contaminated glassware, people's hands or defects of laboratory plumbing. Almost by definition, the most esoteric kinds of cells, requiring the most elaborate mixture of nutrients to make them grow, are those which are the most susceptible to contamination. HeLa cells, the malignant cells derived from a single tumour half a century ago, and widely used as models of malignant cells, are frequent contaminants. Inevitably, when a cell culture that grows only with difficulty is contaminated by cells that grow more vigorously, like HeLa cells, it is a matter only of time, perhaps days, before the original cells disappear, swamped by their more strongly growing competitors. Naturally, people with cells at risk are constantly on the look-out for contamination. This is obviously easiest when the cells at risk can be easily recognized, perhaps because of some distinctive biochemical characteristic. Again, however, and almost by definition, novel types of cells are likely to be least easily recognizable. In other words, it is entirely possible that people may work for weeks or months with cells from a tissue culture without knowing that the originals have been supplanted. It is less easy to see how meaningful results can be wrung from such a culture, but nothing is impossible. The fact that Harris *et al.* have been able to say no more about the fourth of the cell lines than that it is of human origin emphasizes the difficulties of identification. There is therefore no reason to doubt Dr Long's assertion (see page 227) that he believed the cell lines to be authentic; but his response to the doubts raised within and from outside his laboratory seems to have been too casual.

The significance of the paper by Harris *et al.*, for all its interest, should therefore not be exaggerated. Almost certainly, this is not the first occasion on which research reports based on the mistaken identity of tissue culture cells have appeared in the scientific literature. Most of the articles concerned have probably disappeared from people's citation lists as their data have been recognized as spurious or simply puzzling. It would be foolish to ask that the scientific enterprise should be carried on in such a way that such unwanted and unrecognized noise is entirely eliminated. Would intending authors be required to submit with their manuscripts samples of their tissue culture cells for validation by the referees? Would those working with supposedly pure and authentic strains of laboratory mice be expected to follow suit? And how would the scientific journals cope with the problems that would follow? The only reasonable conclusion is that the ideal is probably unattainable, that the scientific literature is based to a very large extent on trust in what its authors say, that the system works surprisingly well and that its occasional failures can be accommodated without too much difficulty. If it follows that the scientific literature is not to be regarded as a collection of Mosaic tablets, nobody will be the loser.

The problems created for laboratories by such developments as those identified by Harris *et al.* are more immediate and in some ways more difficult. The necessarily close working relationship between people sharing the same facilities and even the same objectives in research is one of the most powerful safeguards of the authenticity of the scientific literature. However great may be the temptation for individuals to make too much of preliminary findings or to apply cosmetic treatments to otherwise untidy data, the knowledge that their most candid critics are likely to be their closest colleagues is constantly an influence towards sobriety. It is true, of course, that individuals who have misled themselves can usually succeed in misleading more junior colleagues as well, but even this is an unsure calculation now that many graduate students are zealous custodians of what is right and proper. The more certain safeguard, however, is that people in laboratories where all researchers are in the habit of talking freely with their immediate colleagues about their work, not merely at formal colloquia but over lunch, are imperceptibly persuaded to couch their interpretations of their work in moderate language. (The so-called sociologists of science, always on the look-out for subjects to study, could profitably pay some attention to the sociology of the laboratory as a formative influence on the pattern of science.) No doubt one of the reasons why the most successful laboratories are frequently large laboratories derives from the influence of these local invisible colleges.

Why, then, does this system of polite and imperceptible persuasion by colleagues not always function as it should? In the past few months, the grant-making system has frequently been blamed for departures from the path of strict propriety. The competition for grants, the argument goes, is now so fierce, and the consequences of failing to secure a grant can be so serious for the career of an individual, that people are tempted beyond endurance. This danger is more apparent than real. If the competition for grants is now more fierce, so is the scrutiny of the peer-review committees on whose recommendations most grants are awarded. A more likely explanation is that the grant-making system has disturbed the balance there would otherwise be among colleagues working within the same laboratory. A successful grant-holder tends, when grants are hard to come by, to acquire a status and a degree of autonomy that may not be entirely justified by his colleagues' estimate of his work. And people who are seen to enjoy the goodwill of the grant-making agencies are, if they so choose, to some extent immune from the necessarily searching scrutiny of their immediate colleagues. The Massachusetts General Hospital, one of the best known components of the Harvard medical complex, moved quickly after Dr Long's resignation last year to promise an audit of his work, and the paper by Harris *et al.* is one result. The hospital has probably for the time being done everything it can. But, in the long run, there and elsewhere, the objective should be to make sure that laboratory communities remain as rigorous as they should be in their internal assessment of all their members.

The responsibility lies fairly and squarely with heads of laboratories and heads of university departments. It is an unpalatable responsibility. But unless it is exercised diligently, the result will be unpalatable for the scientific community as a whole. One obvious danger is that of the emergence within the scientific profession of groups of zealous and often over-zealous people dedicated to rooting out what they consider to be dishonesty. Already there are signs of such developments. Some of the allegations of plagiarism in the past year or so appear to have had their roots as much in self-righteousness as in a concern for honesty in the literature. Some of the consequences of these allegations have been unfair to those accused. Yet there is no reason to suppose that the few cases of dishonesty that have come to light are in any sense the tip of an iceberg. On the contrary, as the day to day affairs of reputable scientific journals bear out, there is a general and quite remarkable concern for truth. Would-be authors are forever going back to their benches to check small points raised by referees, who in turn appear almost masochistically prepared to recommend the publication of data or arguments conflicting with their own positions. It would be tragic if these civilized habits were to be corrupted by the activities of self-appointed vigilantes.

More pressures on Rothschild system

Geology project threatened by more cuts

The Rothschild principle, whereby British government departments ("customers") commission research from research councils ("contractors"), already corrupted in medical and agricultural research, is now showing signs of strain elsewhere. The latest development is that two of the customers of the Natural Environment Research Council, the Department of the Environment and the Department of Industry, are threatening to cut the amount of geological research they are prepared to commission.

At stake are the Geological Survey and the mineral reconnaissance programme, both run by the council's Institute of Geological Sciences. The Department of the Environment, which currently contributes £1.5 million to the £4.5 million annual budget of the Geological Survey, is cutting its contribution to geological generally from £3.5 million in 1980–81 to £2.6 million in 1981–82. Although the department has not yet specified its priorities, it is expected that the cuts will be made in the Geological Survey and the engineering geology and bulk minerals resources programmes. The Department of Industry, which is the sole supporter of the mineral reconnaissance programme, is threatening to cut that budget from about £1.2 million in 1980–81 to £0.8 million in 1981–82.

Both departments say that they no longer want to fund research which is not directly relevant to their work. Thus the Department of the Environment will only fund those parts of the Geological Survey related to planning permission. Alternative funds for the survey are unlikely to be forthcoming and it will have to make do with less. Dr G. M. Brown, director of the Institute of Geological Sciences, finds his planning hampered by the Department of the Environment's delay in deciding its new priorities.

The Department of Industry is hoping that the mineral reconnaissance programme will find alternative funds from industry. Since 1972, this unit has been searching for metalliferous ores in Britain, especially in the Scottish Highlands, North Wales and Devon and Cornwall. Although deposits of several minerals — in particular copper and tungsten — have been found, only one, a barites deposit in Scotland, has led to an application for mining rights by private industry. Nevertheless, industry is interested. Companies consider that the results provide useful information for assessing mining profitability in the light of future market trends. But hitherto, there have been no direct commissions from industry.

The Natural Environment Research Council has been trying to establish whether it could act a contractor for private industry. Meetings with the mining houses have thrown up several problems such as that of confidentiality (reports have previously been publicly available) and the assignment of mining rights. Industry itself is keen that the programme should continue, but considers that financial support for it is properly the duty of central government.

Time is now running out. The new financial year begins in April, and twenty of the mineral reconnaissance staff have already moved to oil surveys of the North Sea, paid for by the Department of Energy. The Department of Industry has not finally decided to cut its support, but is thought by several observers to be in a mood to say that if private industry does not produce some cash, it will a sign that the programme has little practical relevance. **Judy Redfearn**

Carter's last budget asks for more

Washington

Determined to enter the history books on an optimistic note, the outgoing Carter Administration has proposed to Congress a budget for the next fiscal year that contains a 4.3 per cent real growth in support for basic research.

"This is the best budget for the past four years for science and engineering research", said Dr Frank Press, the President's Science Advisor and director of the Office of Science and Technology, commenting on the budget proposals last week. He added that it confirmed President Carter's commitment to the support of science and technology "as an investment in the future".

In addition to the intended growth above the expected level of inflation, Dr Press singled out several initiatives that were being proposed by the outgoing Administration "to solve problems that we have known about for a number of years".

One of these is the proposed inclusion in the budget of the National Science Foundation of a $75 million fund to improve university research equipment and laboratories. Their rapidly deteriorating state was highlighted in a recent report from the Association of American universities.

Two other proposals involve traineeships and other awards to avoid potential manpower shortages in fields such as computer science and energy engineering, and efforts to meet the present difficulties of engineering schools where, according to Dr Press, "faculty and equipment are not on a par with what one would expect in industry".

The big question, of course, is how much of this increase will escape the rapidly sharpening budget knife of the incoming Administration of President Ronald Reagan. For many social welfare programmes, the writing is already on the wall, but for science and technology the signals are mixed.

On the one hand, Mr Reagan's budget director, Mr David Stockman, has been talking about the need to rewrite the Carter budget "from top to bottom". Mr Stockman has previously placed the National Aeronautics and Space Administration (NASA) among his "low priority" agencies which might absorb significant cuts. In the same vein, a group known as the National Tax Limitation Committee put out a report last week suggesting, for example, reduced spending on the Galileo mission to Jupiter — already being delayed a year by NASA because of delays with the launch vehicles — on its list of "expenditure control opportunities". The group is also raising questions about the appropriateness of government support for the space shuttle, and suggesting a re-examination of the future cost-benefit ratio.

At the same time, however, members of the various transition teams which have been established by the Reagan Administration to look at research and development programmes have been making optimistic noises, insisting that the new president is committed to the support of "science, technology and productivity".

Part of the increase in defence spending supported by both the Carter and the Reagan Administrations, for example, is likely to have beneficial spin-offs for research support, particularly in areas where the large aerospace companies have a stake. "I have been looking in teacups for the past two months, and each time I look I see a different picture", said outgoing NASA director Dr Robert Frosch last week.

The details of the Carter budget proposals contain several proposed new scientific starts. The NASA budget, for example, contains funding for the development of the Venus Orbiting Imaging Radar — announced by President Carter shortly before the election — as well as a new Geological Applications Program (GAP) which will use remote sensing satellites to study geological resources which might contribute to the discovery of new oil and gas deposits.

Reflecting the delays in the space shuttle programme, the proposed commitment to start work on a fifth shuttle orbiter, the

0028-0836/81/040213-02$01.00

additional costs of science projects such as the space telescope, and the new projects described above, the total requested budget for NASA comes to $6,700 million. This would be a 20 per cent increase over the budget for the fiscal year 1981 which began last October, and if it is allowed to stand, would be the largest annual increase in the agency's budget since the early 1960s.

The National Science Foundation has also put in for a hefty 23.5 per cent increase, from $1,096 million in the current year to $1,353.5 million next year. Most of this reflects the Carter Administration's keenness to support both research and training in engineering fields. The new engineering directorate will receive a 20 per cent increase in its research budgets.

At the National Institutes of Health (NIH), the proposed increase for biomedical research is less spectacular. On the basis that whatever the president asks for is traditionally increased by Congress, Mr Carter is suggesting that the NIH budget for basic research be raised by 9.4 per cent. Allowing for inflation, this would result in a drop of 1.1 per cent between 1981 and 1982.

At the Department of Energy, increased support for research into synthetic fuels and nuclear power — particularly fusion energy — has resulted in a requested increase of 9.4 per cent in real terms for basic research, second only to that of NASA.

Many of these figures will remain only as indications of the "good intentions" with which the Carter Administration is leaving office. Perhaps of more lasting significance are the figures prepared by Dr Press to demonstrate the main trends of federal support for science during Mr Carter's four years in the White House.

These reveal, for example, that overall the biggest winner as far as support for basic research is concerned has been the Department of Defense. If the proposed 1982 budget figures are taken into account, the Pentagon's basic research efforts will have grown by almost a quarter — 22.6 per cent — between 1978 and 1982.

Next come NIH. Congressional enthusiasm has raised the NIH research budget by 13.3 per cent over the four years, compared with a growth of 12.8 per cent at the Department of Energy.

The National Science Foundation, even if it is granted this year's large increase, will still only have seen its basic research grow by 9.2 per cent. And at NASA, reflecting the pressures which the space shuttle has imposed on the space science programmes, the basic research budget actually fell, in real terms, by 0.6 per cent over the same four years.

Overall, the growth in basic research, including the 1982 proposals, would come to 10.8 per cent for the period of the Carter Administration. In current dollars, the budget would grow by 58.2 per cent. from $3,704 million to $5,801 million.

Ironically, the research and development budget shows an identical increase of 58.2 per cent from $26,388 million to $41,734 million.

"Anyone who says that we do not engage in long-term planning is proved wrong by these figures", quipped Dr Press — expected soon to be elected president of the National Academy of Sciences — although he added that the agreement was actually fortuitous and that "the figures just happened to fall our this way". In practice, he will not be required to explain why it should be otherwise.

David Dickson

Argentinian power

Soviets help

Argentina has bought five tonnes of heavy water from the Soviet Union under International Atomic Energy Agency safeguards, the Argentinian Comisión Nacional de Energia Atómica announced last week. It is intended for "topping up" the Atucha-1 nuclear station, which needs on the average an annual heavy-water replacement of 1.5 tonnes.

The sale is part of the growth of Soviet–Argentinian trade since January 1980, when Argentina refused to back President Carter's embargo of grain sales to the Soviet Union. The Soviet Union is now Argentina's main market for agricultural products. In July of last year, Argentina's Secretary of Commerce, Alejandro Estrado, signed an agreement to supply the Soviet Union with 20 million tonnes of feed grain and soya beans during the next five years. There are also persistent rumours that a major agreement to export meat to the Soviet Union is now being negotiated.

In return, the Soviet side has shown considerable interest in Argentina's nuclear programme. Argentina has been conspicuous among third world countries since the early 1950s for its nuclear programme aimed at ultimate autonomy in both research and technology. The commission has announced that the target will be for practical purposes attained by the end of 1981, when the Córdoba uranium processing plant will begin producing an estimated annual production of 150 tonnes. Rafael Coppa, the director of the plant, said last year that Argentina will then have full control of the primary uranium cycle, from prospecting for

---

Promises for President Reagan to deny

Specific proposals included in the budget are as follows.

- Major difficulties with the development of a vehicle to launch the two Galileo spacecraft on their journey to Jupiter from the space shuttle have caused NASA to propose delaying the Galileo project for one year and switching to a new launch vehicle, a converted Centaur rocket.
- The Carter Administration proposes that the National Institutes of Health should aim to stabilize support for both competitive research grants and research traineeships. Last year, the Administration promised to provide enough money to keep the number of new and renewing competitive research grants constant at about 5,000. Given general fiscal constraints, however, this meant cutting back severely on the number of traineeships, a move which brought strong protests from various sectors of NIH, and was subsequently overturned by Congress.
- Defense Department support for research on US campuses seems destined to continue to grow faster than support from any other federal agency. According to the budget request for the fiscal year 1982, military funding for research and development at US universities and colleges will be 21 per cent greater than in 1981, totalling $639 million.

The proposed figure, most of which will be spent on unclassified basic research projects, is part of a 16 per cent increase in all military-sponsored basic research expenditure.

- One major new start proposed by the National Science Foundation (NSF) is the detailed design and initial construction of a 25-metre millimetre-wavelength radio-telescope which is planned for installation at Mauna Kea in Hawaii.

Plans for the new telescope have generated widespread support in the radioastronomy community. NSF has asked Congress for funds as part of a 29 per cent increase in the NSF budget for astronomical sciences, rising from $58.5 million to $75.6 million.

- The Carter Administration is proposing to cut support for research and development in the Environmental Protection Agency (EPA) from $364 million to $345 million. At the same time, the outgoing Administration wants to earmark an extra $28 million to improve its review of the environmental impact of proposed major energy projects in the west of the country — particularly in connection with the synthetic fuels programme — and to launch a government-wide research programme on the effects of acid rain.
- The Carter Administration wants to give a major boost to research into magnetic fusion. The 1982 budget proposals include an increase of 28 per cent in fusion research, to a total of $520 million in budget obligations. $32.8 million of this would be spent on a new centre for magnetic fusion energy, as proposed by the Administration following a thorough review of the magnetic fusion programme last year. Research on magnetic confinement systems would increase from $119 million to $151 million.

David Dickson

0028-0836/81/040214-02$01.00

uranium to putting the fuel elements into the reactor.

Nevertheless, the search for self-sufficiency does not rule out cooperative projects. Indeed, on returning from the World Energy Conference in Munich last September, Admiral Carlos Castro Madero, the president of the Argentinian commission, openly attacked the "negative aspects" of restrictions of nuclear technology transfer, adopted first by the "London Club" of nuclear suppliers and then (at the end of 1979) by a wider circle of industrialized Western states.

Soviet interest has grown in the past year. Last April, during a visit to Buenos Aires, the Soviet foreign trade vice-minister Aleksandr Manzhelo suggested a major nuclear cooperation between the two countries, stating at the same time that he thought that Soviet–Argentinian trade could well double in the next few years. At the end of July, Yurii Fokin, Secretary of the Soviet Foreign Ministry, visited the Atucha-1 plant. **Vera Rich**

Large electron project

Swedish cloud

Sweden is having second thoughts about participating in LEP, the 500-cm 500-GeV electron–positron colliding machine which, at a cost of 900 million Swiss francs, is planned to be the next major project of the European centre for high energy physics research, CERN. At a meeting of the Swedish Natural Sciences Research Council last month, 90 per cent of those present expressed doubts about the arrangements for funding the project.

Delegates from CERN's member states are expected formally to approve the building of LEP at the next meeting of CERN council in June. The plan is to finance the project out of CERN's annual budget by reducing expenditure on other programmes such as the intersecting storage rings and the synchrocyclotron. How quickly LEP can be built will depend on how much of the budget — SwFr610 million this year — can be diverted to it each year. What seems to be worrying Sweden is that the CERN council, which requires a two-thirds majority vote to approve budgets, could demand that Sweden pay more if the cost of LEP rises above initial estimates.

One faction of the Swedish research council says that LEP is simply too expensive to be built now. Another would agree to the project with some concessions — either that Sweden be made exempt from budget increases approved by CERN council, or that the CERN budget be divided into LEP and other programmes, giving Sweden the option of leaving LEP while remaining a full member of CERN's other activities. Under the present arrangement, a decision not to participate in LEP would effectively be a decision to opt out of CERN altogether, An incidental factor which seems to have added weight to the arguments of LEP's Swedish opponents is the feeling that Swiss industry has reaped unfair benefit from contracts arising out of CERN's work.

Sweden's objections, which have come rather late in the negotiations (most of CERN's members have informally agreed to the LEP proposal), could delay official approval for the project. Although Sweden is not alone in wanting some guarantee that costs will not get out of hand, making it exempt from cost increases approved by CERN council is unlikely to be popular with other members: neither is splitting LEP from the rest of CERN's budget. Many see incorporating LEP into the annual budget as a way of controlling the rate at which money is spent on it.

At present Sweden contributes slightly less than 4.3 per cent of CERN's annual budget. If it decided to leave LEP, the other member states would have to decide how to redistribute the costs amongst themselves or whether to extend the time taken to build the machine. The chairman of the Natural Sciences Research Council, Mr Mats Lemner, expects that the council will shortly discuss the contribution with the Minister of Education. Depending on the outcome of this discussion, parliament may have to make the final decision. If it does, the Swedes would to be able to meet CERN's June deadline for a decision on the contribution. **Judy Redfearn**

Artificial hormones

European register?

Brussels

The European Commission is making heavy weather of its plan to ban the use of certain artificial hormones in animal farming. A meeting of agricultural ministers planned for last week was cancelled after the death of Mr F.O. Gundelach, the Danish agricultural commissioner. But the signs were that the meeting would have failed to reach an agreement.

The Commission decided last September that something should be done about hormones after the discovery that diethylstilboestrol was still being used for veal production in Italy. The hormone is banned in the United States and also in many European countries, but is so effective at increasing weight-gain in calves that, where its use is banned, black markets such as that in Belgium spring up.

The agricultural ministers were to have discussed two proposals elaborating on an original proposal made last December. The first calls for a register to keep track of all hormones used as medicinal products, whether for human or animal use, from manufacture and storage to distribution and final use. Veterinarians would be required to control all administration of hormones to animals. The proposed rules say that banned hormones can be used only for "therapeutic treatment" of pathological cases diagnosed by veterinarians, and not for chronic use in preventive treatments. The second proposal is that there should be a comprehensive sampling system for testing animals and carcasses

The proposals as drafted involve the "positive listing" of those hormones considered safe for use in meat production. This is opposed by Belgium and the United Kingdom. There is no dispute over banning compounds such as diethylstilboestrol, but it is held that a positive list would inhibit the development of new materials.

A system of "negative listing" would overcome some of these problems, but this would have to be continually revised as alternative hormones came on the market. Given the political need for action, the Commission favours the more cautious positive listing system. Whichever route is followed, the cost to European farmers of making the necessary adjustments will be substantial. **Jasper Becker**

Soviet chemical industry

Effective economy

The Soviet chemical industry is rapidly acquiring prestige status in the Soviet media and ranks, according to a *Pravda* article last week, together with nuclear energy, space research and electronics, as one of the hallmarks of twentieth century progress. The proximate source of this accolade is not hard to identify: almost every article cites, at some point, Mr Brezhnev's dictum that "there can be no effective economy today without a modern large-scale chemical industry".

When Mr Brezhnev made this pronouncement at last October's plenum of the Central Committee of the CPSU, however, he was not so much commending the industry as calling for a programme of "resolute measures" for overcoming major shortfalls in chemical production, ranging from chemical fertilizers and plant protection agents to synthetic fibres, dyes and household detergents. The 33 per cent production increase specified in the guidelines for the new Five Year Plan for the chemical and petrochemical industries (recently placed under separate ministries) is, say the planners, essential if the shortfall is to be eliminated.

The chemical industry does not shoulder full responsibility for the gap. At the end of December, a *Pravda* editorial shifted at least part of the blame to other sectors. Fertilizer plants, said *Pravda*, were held up by insufficient supplies of natural gas and "inaccurate" planning by the light metallurgy sector. Plants with processes requiring high temperatures and pressures often cannot obtain corrosion-proof equipment. In some cases new factories have been built, without the necessary equipment being forthcoming, while, on other occasions, expensive installations have been purchased before it has been

0028-0836/81/040215-02$01.00

finally decided just what the plant is to manufacture. The state supply board "Gossnab", *Pravda* concludes, must supervise the whole supply process more closely to eliminate such discrepancies.

Far more serious, however, are suggestions raised last week by a group of scientists from Byelorussia that the implementation of radical new technologies is being blocked by interdepartmental wrangles. In a major *Pravda* article they describe two such cases, in sectors which all planners consider to be of the highest priority — energy and agriculture. The first project, highly thought of by the Institute of High Temperatures of the Soviet Academy of Sciences, would have combined an electrolysis unit with a nuclear power plant so that hydrogen for fuel or industry could be produced during "off-peak" times. This combination of chemistry, electrochemistry and nuclear power engineering ran into the stumbling block of what the group describe as the "excessively rigid specialization of the branches of industry".

The second project, for the production of protein biomass for animal feed using hydrogen-consuming bacteria, was worked out jointly by teams from the Byelorussian and Moldavian academies as long ago as 1976. To be cost effective, however, the byproducts of the synthesis — oxygen, nitrogen and possibly ammonia — would also have been exploited. All these are prime requisites of the mineral fertilizer industry, which itself produces waste gas with a usable hydrogen content. Yet although plans have been drawn up for a pilot biomass plant to be run in conjunction with the "Azot" fertilizer plant in Grodno, neither the Ministry of Chemical Fertilizer Production nor *Glavmikrobioprom* — the body in charge of microbiological production — can see its way to go ahead. **Vera Rich**

European research

Counting costs

Brussels

The knotty problem of putting a value on the European Community's research and development programmes is the subject of new proposals which the European Commission has just put forward. The objective of this soul-searching is to define criteria and methods for both evaluating and exploiting research. The idea stems not from a fit of self-doubt but from the Council of "Research" Ministers in December 1979. There are no hard and fast ideas in the document, but areas of investigation are instead defined. Four pilot projects have been approved, which will give independent experts free rein to put certain programmes under scrutiny. The commission proposes then to use the experience of the experts to formulate a policy for exploiting and evaluating research results. This is to be ready by the end of the year but, before then, the first outlines of the policy will be presented at the International Congress for Research Evaluation in October.

One pilot project has already been completed by six independent experts, who looked into the Community's energy conservation and solar energy activities. Four other expert groups are to investigate programmes of research and development in fusion, radioactive waste management, the "Reference Bureau Programme" concerned with physical standards and critical data, and geothermal energy, the use of hydrogen as a fuel and work on the systems analysis of energy use.

The tightening of the European Community's budgetary belt is one motive for the new initiative, but another is the fear that the results of research are being ineffectively disseminated. The experts will be concentrating on the diffusion of information and in particular how to make the results available to the layman, requiring that more outlets should be found and language barriers overcome. Indeed, one member of the European Parliament has asked in a written question why the Commission does not publish the results of research in English as well as in the author's mother tongue. But the commission says that unless the results of research are first digested, readership is always restricted to specialist circles.

The stickiest part of the process on which the Community has embarked is the evaluation of research whose impact is likely to be long term or even unforeseen. Simplifying the results may help. Thus the report on solar energy and conservation says that the success of the programme cannot be measured in "tonnes of oil saved" and that the use made of research carried out may require further action and will take time. Even so, the chairman of the evaluation team, Ugo Forinelli, is confident that European research in this area is already competitive with that in the United States even though the European budget is much smaller than the American. Forinelli's team has made a number of recommendations about the management of the Community's programme, including the need for more explicit statements of objectives when putting contracts out to tender and some suggestions of areas where research might be concentrated or even discontinued.

The commission seems modestly aware of the inherent limitations of plans to evaluate research. The underlying objective is to strengthen the link between research and development and industrial innovation, and the commission acknowledges that its own research cannot be an important source of commercial inventions. But restrictions on Community expenditure outside the agricultural field are likely increasingly to stimulate the justifying of research. **Jasper Becker**

Swedish research

Lucky science

Stockholm

Research is one of the very few areas to be given more money in the budget bill for the fiscal year 1981–82, presented this week in Stockholm. Against the background of a budget deficit amounting to about $15,000 million, cutbacks in social services and mounting economic gloom, research received an extra $28 million.

Considering that this money is to be divided between the projects of eight ministries, individual increases will not be very great, and in real terms may well be eaten up by inflation, which was 14 per cent in 1980. One staff member at the Natural Sciences Research Council estimated that the net results of the council's $2 million increase will be an unchanged level of activity. But this is better than a cutback, the fate of nearly all other sectors.

The government sees research strengthening Sweden's industrial competitiveness in the long term. The Minister of Industry, Nils Aasling, wants to develop sophisticated chemical and mechanical products, and is to make specific allocations for this purpose in the spring. One sign of this policy is the changing proportions of the space budget being allocated between international and national programmes. In the 1980–81 fiscal year, the Swedish Space Corporation's budget of about $44 million was equally divided between international and national activities, but in future the accent will be on national programmes. The corporation's budget has been increased overall to about $58 million — a real increase in spite of inflation — about half of this will be used to build the country's first space satellite, Viking, which is to be launched from the European rocket Ariane in May 1984. It is hoped to produce the Viking at half the cost of comparable satellites from other European countries, and the project is planned to give Sweden a profitable national space industry. Another project being planned is an experimental telecommunications satellite.

The lion's share of the increase in research expenditure — about $16 million — is to be disbursed by the Education Minister, Jan-Erik Wickstroem, who wants to increase research capacity across the board. He is proposing an increase of about 12 per cent for mathematics and natural science faculties, an average of about 9 per cent for the four research councils as well as 90 new PhD places to be spread over all subjects, $2 million for particularly expensive equipment, $2 million to subsidize lecturers who want to take time off for research and $1 million for research libraries. Fifteen new chairs will be set up, among them molecular genetics (Karolinska Institute), medicine (Uppsala) and astrophysics (Gothenburg).

Wendy Barnaby

0028-0836/81/040216-01$01.00

Graduate education

Too many laggards

The Science Research Council, the chief source of public support for British graduate students, is alarmed at the lengthening time taken to complete PhD courses. A preliminary survey of 25 higher education institutions has shown that on average only 60 per cent of those holding SRC studentships complete their PhDs within four years. According to Sir Geoffrey Allen, chairman of the council, a figure of 80 per cent would be respectable, but 90 per cent would be the ideal.

The issue has come to the surface after an investigation by a working party of the Advisory Board for the Research Councils, which has been looking into the broader question of postgraduate education and manpower needs. That in turn was stimulated by the revelation of Sir Michael Posner, chairman of the Social Science Research Council, to the Public Accounts Committee of the House of Commons last summer that fewer than 30 per cent of his council's graduate students complete their PhDs within four years.

Given pause by that statistic, the Science Research Council made a rapid survey of five institutions where it supports students and arrived at the figure of 60 per cent, which a subsequent survey of 25 institutions has upheld. The advisory board's working party, which is also concerned with the performance of postgraduates supported by the Social Science, Medical and Agricultural Research Councils, has commissioned a more detailed study intended to throw some light on why so many students take so long to complete their theses, or even fail to complete them at all.

Particular attention is likely to be paid to the performance of the Science Research Council's students, if only because there are more of them than of the other councils — on the average, 2,350 new science studentships are awarded each year. Most of these are in the gift of university departments, to which studentships are allocated on a quota basis. Studentships are worth about £3,500 a year, and are tenable for three years, the estimated time for completing a research project.

The reasons for these delays are still obscure. The Science Research Council expects to find marked differences of performance in different institutions and subject areas. PhDs in pure science may more often be completed than those in applied science — applied scientists and engineers are more likely to find jobs in industry, where writing a thesis may seem irrelevant and where there is little time for writing up anyway. It is also suspected that institutions and departments with a large number of PhD students will have a better track record than those with relatively few.

The issue also, however, raises questions concerning the meaning and purpose of a PhD, which the advisory board's working party is looking into under its broader remit. Should a PhD for example, be a thorough and lengthy investigation of a detailed scientific problem, or more simply a means of training a student in the techniques of research? The approach is bound to have implications for the completion time.

As yet, no reliable pecking order of institutions has been established, but the early surveys do suggest that the Universities of Birmingham, Cambridge, East Anglia and Bristol, and King's College in the University of London, have the best completion records and that the Universities of Newcastle upon Tyne, Sussex and Bradford, together with Imperial College, London, and most of the polytechnics, have the worst.

The poor track record of Imperial College, regarded as a highly prestigious scientific and technological institution, may seem surprising. Lord Flowers, rector of the college, says that the explanation may be that large numbers of its postgraduates rapidly find employment in industry, leaving them little time for writing up.

The advisory board's working party has yet to decide what should be done. Sanctions against departments with poor track records have been mentioned. Cutting quotas of studentships is an obvious device. Sir Geoffrey Allen, however, hopes to avoid such heavy-handed treatment. Most academics, he says, are willing to accept genuine criticism and put their houses in order. The peer review system should take care of that.

Judy Redfearn

British universities

More confusion

Confusion among British universities about their financial prospects appears to have been further deepened by the letter from the chairman of the University Grants Committee, Dr E. S. Parkes, circulated to vice-chancellors on 30 December. The letter contained a warning that the resources available for the 1981–82 academic year may be reduced by between 5½ and 6 per cent compared with the amounts advertised in the Public Expenditure estimates a year ago. The committee's latest estimate of the shortfall next year is an amalgam of a 3½ per cent cut estimated to be the universities' share of the £30 million cut for higher education announced last November and the still incalculable effect on university finances of the partial disappearance of overseas students, some of whom have been frightened away by "economic" fees.

Some universities regard Dr Parkes's warning as a signal for drastic belt-tightening. Last week, for example, Lord Annan, Vice-Chancellor of the University of London, told the university senate that the total budget of £200 million might be reduced by between £15 and £20 million in 1982. Other universities appear to be taking a more phlegmatic line, believing that they cannot know the worst until the individual allocations of funds for 1981–82 are made in April or soon thereafter.

Between now and then, the committee itself will have several difficult questions to decide. One possibility raised in the letter from Dr Parkes is that the committee may keep back until later in the academic year a proportion of the funds made available by the Department of Education and Science, using the reserve to make good deficiencies that have by then appeared. There are precedents for such a reserve, but no decision has yet been made, nor have criteria for deciding how to use the money been defined.

The Parkes letter also promises explicit guidance in the spring on the numbers of home students at which British universities should aim in the coming academic year, and on their desired distribution among different kinds of courses. Although such "guidance" has accompanied previous financial allocations to British universities, some universities now apparently fear that the committee intends to be more *"dirigiste"* than in the past — while others remark that Dr Parkes's letter is a good deal less so than his speech to a closed meeting of the Committee of Vice-Chancellors at the beginning of December. It does however seem clear that the University Grants Committee will take steps to ensure that universities do not earn their way out of trouble by recruiting more home students than at present for the sake of the extra fee income that would bring.

Planning for the year ahead has been further confused by reports that Mr Mark Carlisle, Secretary of State for Education and Science, told a conference in the north of England on 6 January that he saw no reason why the cuts now proposed should be matched by "a reduced provision". Universities, on the other hand, say that there will have to be a reduction of student entry this October if they are to live within their straitened budgets.

The future constitution of the University Grants Committee itself also appears to be in question. Taking its cue from a recommendation of the House of Commons Select Committee on Education towards the end of last year that the committee should cultivate more independence, the department has raised the possibility that the committee's staff (at present seconded from the Civil Service) should become direct employees. Some of those concerned wryly reflect that this is not quite the independence that the House of Commons committee had in mind, but they acknowledge that such a step would reduce the size of the Civil Service and transfer part of its cost away from central government. The fine print in the proposal, which is said to be "very small", is being read carefully.

0028-0836/81/040217-01$01.00

CORRESPONDENCE

Badgers guilty

SIR — In his letter published on 11 December 1980 (page 532), Stephen Harris, among many other misleading and biased remarks, stated that the rate of decline in the incidence of tuberculosis (TB) in badgers in the South West was "paralleled" by a decline in the incidence of reactors in cattle herds not only in the South West but also in the rest of England. This statement was based on a graph which was attached to a copy of the letter sent to Lord Zuckerman and which is now reproduced (Fig.1). It was deduced from this that there might well be some common cause unconnected with the gassing campaign that was reducing the incidence of TB both in cattle and badgers throughout the country.

This "parallelism", however, was only secured by expressing all plotted values as percentages of the corresponding 1974 values. In Fig.2 I have plotted the breakdowns for the South West and for the rest of England on the same scale, subdividing the South West into (*a*) Gloucester, Avon and Wiltshire; (*b*) Cornwall; and (*c*) Devon, Dorset and Somerset. The radical differences between the counties that have a high incidence of herd breakdowns, (*a*) and (*b*), and the rest of England (*d*) is immediately apparent.

Table 10 of the report (*Badgers, Cattle and Tuberculosis,* Lord Zuckerman; HMSO, London, 1980) shows that the sources of infection are indeed different in the South West and in the rest of England. In groups (*a*) and (*c*) three-quarters of the infections are attributed to badgers, but only 1 per cent to imported Irish cattle, whereas in the rest of Great Britain one half are attributed to imported Irish cattle and none to badgers.

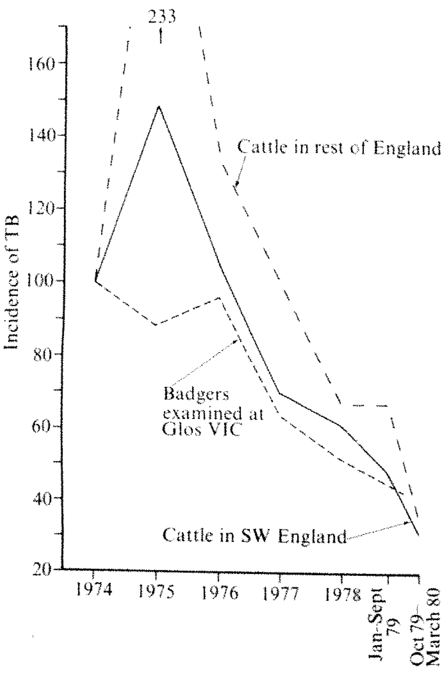


Fig.1 Annual changes in the incidence of TB in badgers sampled at the Gloucester Veterinary Investigation Centre (page 62 of the report) and herds of cattle in South-West England and herds of cattle in the rest of England (page 64). The first year for which data were available in the report from all three samples (1974) was taken as 100 per cent; for subsequent years, the level of TB was expressed as a percentage of the incidence in 1974.

In Cornwall (*b*), as Harris observes, only 15 per cent were definitely attributed to badgers, but as almost three-quarters of the causes of Cornish breakdown were recorded as "unknown" and the remainder were either from purchased cattle or contiguous premises, it may be presumed that infection from badgers was much greater than this.

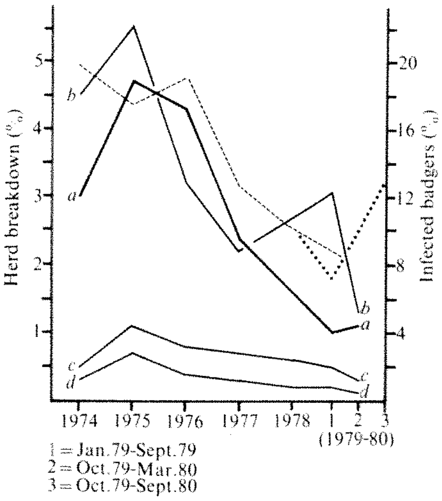


Fig.2 The percentages of herd breakdown are shown by full lines (from Table 19 of the report). The percentages of badgers found to be infected are shown by the broken line (Table 15). The dotted line gives some supplementary information from Table 18 (updated).

The fall in the incidence of herd breakdowns following the introduction of the gassing campaign in Gloucester/Avon agrees well with the reduction in the incidence of TB in badgers (broken line), bearing in mind that in addition to the reduction in the percentage incidence in badgers the numbers of badgers in close contact with herds must have been substantially reduced by the gassing.

The dotted line is derived from the latter part of Table 18, updated to September 1980, and indicates the rise in TB incidence in badgers following the cessation of gassing. (Updated figures for herd breakdowns in Gloucester/Avon and Wiltshire are not yet available).

The only support that Fig.2 gives to the suggestion that there was a common cause, other than appropriate action by ministry officials, affecting the decline in TB operating throughout the country is the substantial rise between 1974 and 1975. This cannot be related to infection from badgers. It may possibly have been due to climatic factors or to some improvement in the tuberculin test.

Figure 1 is a sad illustration of the way in which distorted and partial graphical presentation can be used for propaganda purposes.

FRANK YATES
Rothamsted Experimental Station,
Harpenden, UK

Rothschild retreat

SIR — It is not clear from your leading article in the January 1/8 issue (page 2) whether you regret the government's "retreat from Rothschild" or just the manner in which it is happening.

One of the difficulties for the Medical Research Council (MRC) at the time of the Rothschild report was the uncertainty as to the failures in the then current system which it was intended to correct. The Ministry of Health did not identify any omissions on the part of the MRC in their contribution to improving the health of the nation, nor did they subsequently propose contracts in fields of work outside the previous commitment of the MRC. Rothschild, in a thoughtless comment, referred to the very low expenditure on research into ageing, one of the fundamental aspects of all living things, but the understanding of which is nowhere in sight. He failed to ask the practical question as to how much of the MRC budget was spent on investigating diseases most common in old age. The answer would have been a major part.

There are, at present, three developments in medical research on which great hopes of practical gain are based: interferon, antibody production by hybridomas and DNA technology. How far the expectations will be justified remains to be seen, but the heavy investments of pharmaceutical companies throughout the world in these fields is evidence that the hopes are held widely. MRC laboratories were responsible for the first two discoveries and made a very big contribution to the third. None came from planned research but from "Dr Gowans' doctrine of investment in good ideas and good people".

It has been argued strongly that much more attention should be given to public health and preventive medicine and that the MRC's expenditure has the wrong emphasis. But one of the major advances in this field, the correlation of the incidence of lung cancer with cigarette smoking, came from work in an MRC unit.

No doubt governmental administration of civil science could be improved substantially, but the next enquiry which you appear to be advocating must surely start by identifying the omissions and failures of the present system whether in medical, agricultural or any other field of research before recommending changes to overcome them. Your generalized denigration of a system which has achieved obvious successes is not enough.

R.R. PORTER
Department of Biochemistry,
University of Oxford, Oxford, UK

Creating problems

SIR — Your readers should know of the court action by creationists against the State of California, now expected to begin on 2 March. The suit, which is sponsored by the Creation Research Society of San Diego, is in the names of Segraves *et al.*, students in California schools, including children of Mrs Nell Segraves of the society. The suit charges that the State of California has violated the

Continued on page 335

0028-0836/81/040218-02$01.00

Deep-ocean hydrothermal vent communities

from J.T. Enright, W.A. Newman, R.R. Hessler & J.A. McGowan

THE theory of plate tectonics has led to a completely unexpected bonanza of research problems for marine biologists. The theory predicted major heat input to the ocean in areas where the sea floor is spreading, and in the search for those sources, geophysicists discovered not only deep-ocean hydrothermal vents, where warm water emerges from subterranean sources, but also extraordinary concentrations of large and unusual animals surrounding the vents[1,2]. The bottom of the deep sea typically has a sparse fauna dominated by tiny worms and crustaceans, with an even sparser distribution of macrofauna, including clams, crabs and echinoderms. Near the hydrothermal vents, however, are bizarre ensembles of abundant large animals — communities which resemble nothing ever before encountered by marine biologists. The most conspicuous and unusual animals around the warm-water vents (∽ 20°C) are huge gutless tube-worms (Vestimentifera), some of which are over 3m long and up to 10 cm in circumference; but there are also remarkable densities of huge clams (25cm length is typical), blind crabs, fish and polychaetes (see figure). More than 100 species have been collected[2], many of which are elsewhere unknown.

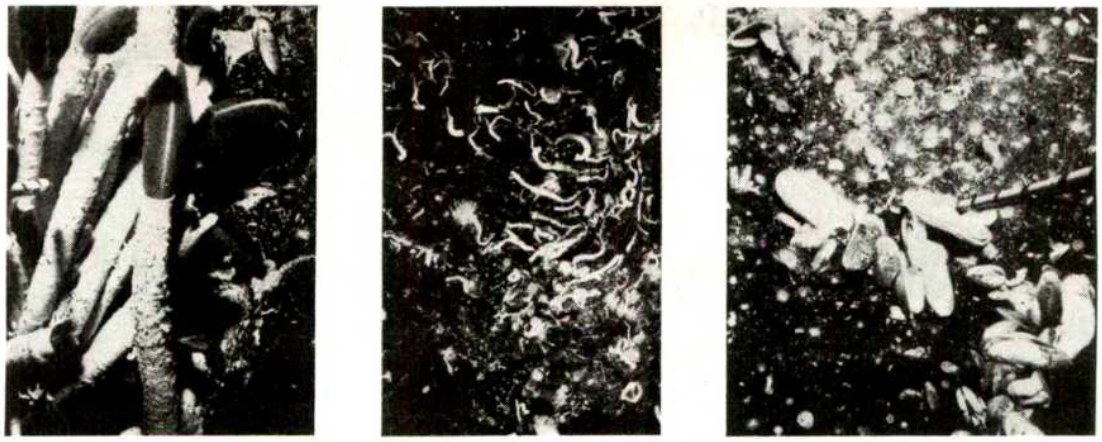
How do the animals in these deep-sea communities find enough to eat? Elsewhere, the deep benthic fauna relies on particulate matter, ultimately derived from photosynthesis, falling from above. The food supplies necessary to sustain the vent communities, however, must be many times that ordinary fallout. In the first report which described the vent fauna, Lonsdale[1] proposed two possible sources of nutrition: advection of food materials from surrounding regions and bacterial chemosynthesis. Shortly thereafter three kinds of evidence were presented[3] which support the idea of intense local chemosynthesis: (1) hydrogen sulphide, as a potential basis for fixation of CO_2, was found in vent water; (2) some of the bacteria isolated from a vent site were found to be capable of chemosynthesis on a thiosulphate medium; and (3) bacterial concentrations of 10^8 to 10^9 cells per ml were found in samples thought to be pure vent water. This final observation seems decisive: if such astonishing concentrations of bacteria — difficult to achieve even under optimal laboratory conditions (D.M. Karl, personal communication) — are indeed typical of steady-state vent outflow, then food from within the vent would be dwarfed by any contribution from advection. Hence, the widely quoted conclusion was reached, that bacterial chemosynthesis provides the foundation for hydrothermal-vent food chains[2-5] — an exciting prospect because nowhere else on Earth is a major community of diverse organisms known which does not depend, at least indirectly, on photosynthesis.

There are, however, certain difficulties with this interpretation. Some of the large sedentary organisms associated with vents are found at ordinary deep-sea temperatures many metres from the nearest hydrothermal sources[6], and therefore would not directly encounter any bacteria from within the vent. It is conceivable, however, that bacterial aggregates, which have grown subterraneanly, in the rising plume of vent water might rain back out in the peripheral areas to nourish these filter-feeders. Another difficulty is that similarly dense populations of large animals have been found[2,7] in the proximity of 'smokers' — hot-water vents where the dominant, conspicuous flow emerges at temperatures up to 350°C. No bacteria can be expected to multiply or even survive such heat, and none was found in the hot vent water[7]. Unless the smokers are consistently associated with smaller but more hospitable warm-water vents[2] — a distribution that has not actually been observed (F.N. Speiss, personal communication) — the chemosynthesis hypothesis could at best account for only a fraction of the vent faunas.

The food resource question has been further complicated by a recent and more extensive investigation[5] of the microorganisms associated with one of the warm-water vents in the Galápagos Rift region. Bacterial concentrations were measured in near-vent water which were approximately three orders of magnitude lower than those initially reported (5×10^5–10^6 cells per ml). These values were confirmed by replicate estimates of biomass based on ATP concentrations in the water[5] (100–250 µg of microbial carbon per litre, compared with the previous estimate[3] of 100–1,000 mg of bacteria per litre). Although the authors of that report still endorsed the chemosynthesis hypothesis[5], the available data suggest that Lonsdale's first alternative[1] deserves more attention than it has received. Could advection conceivably supply the near-vent fauna with enough food organic matter ultimately derived from photosynthesis?

Because of its temperature, water from a vent is very buoyant, densities as low as 0.62 gm cm^{-3} having been reported for a 'black smoker'[8]. The rising plumes of vent water are, however, rapidly diluted. Spiess *et al.*[7] report that a vent discharging at 23°C into water at 1.8°C produced a temperature only 0.3°C warmer than ambient at a height of 10 m above the vent. These values indicate that within 10 m of the bottom, the advected diluting water was about 70-fold the discharge volume $[(\Delta T_1-\Delta T_2)/\Delta T_2]$. Stability considerations dictate that such advective flow into the vent area would originate near the water–substratum interface, where suspended organic matter accumulates and epibenthic animals are concentrated. Typical values for suspended particulate matter in deep-sea waters within a few metres of the bottom[9,10] are of the order of 50–500 mg m^{-3}, of which several per cent is organic matter. Taking a conservative estimate of 5 mg of particulate organic carbon per cubic metre and the calculated dilution factor of 70, we find that for every cubic metre of vent discharge, 350 mg of particulate organic carbon would be advected into the vent area. If discharge volume were to be, say, 1 m^3 s^{-1}, advection could provide more than 30 kg per day of potential food. In addition, there is a good chance that small live animals in the advected water might be killed or stunned by thermal and/or

0028-0836/81/040219-03$01.00



Examples of hydrothermal-vent fauna from Galápagos Rift area. Left, vestimentiferan worms, about 3 cm in diameter. Middle, dense group of encrusting serpulid polychaetes, about 5 cm long, some with tentacular crown extended. Right, vesicomyid (white) and mytilid (grayish) bivalves adjacent to an aggregation of anemones. The tip of the probe is 1 cm in diameter.

chemical shock, and subsequently rain down, thereby contributing to the vent food supply. Hence, it appears that an input of non-photosynthetic food might not, at least in principle, be required by the vent animals.

Such input calculations, however, ignore what may be the most important aspect of advective processes, that the near-vent area should accumulate and concentrate most advected organic matter. Entrained food particles which are not consumed in their initial passage through the vent area would be wafted upwards, but ought thereafter (given a density greater than ambient seawater) to rain out predominantly on the near-vent substratum. Once on the bottom, this food material would not be lost to the vent ecosystem because of resuspension. Sediment-trap data[11] indicate that the particulates suspended in near-bottom water of the deep sea must, on the average, be resuspended every few days. Because of such resuspension, the initially advected material would be subject to repetitive re-entrainment, thereby enriching the advected waters. Hence, filter-feeding animals near a vent should experience not only increased flow rate of waters bearing organic nutrients, as suggested by Lonsdale[1], but also experience unusually high concentrations of suspended particulates from non-vent sources, food which is ultimately derived from sea-surface photosynthesis. The nutrition of filter-feeding animals is even more dependent on food concentrations than on the flow rate of surrounding waters. At any new vent site, even if chemosynthesis were to be negligible, this sort of accumulation should progress until catabolic degradation, by both micro- and macroorganisms, balances the net rate of advective input.

Both Lonsdale's suggestion that food materials would be advected from extensive areas surrounding deep-water hydrothermal vents, and the enhancement of the local density of these food materials due to gravitational fallout and recirculation, appear to be reasonable, if not logically inescapable, expectations. The as yet unresolved issue is the quantitative one — what proportion of the food supply for macroorganisms is provided by such advective input and what proportion by vent-related chemosynthesis? The suggestion that advective input alone might provide an adequate nutritional basis for near-vent food chains implies that animals relying on advected biomass ought to do well around any intense persistent deep-water source of heat — an idea which may be worthy of direct experimental test.

By similar reasoning, the near-vent area should also represent a zone of accumulation for any organic materials which might be produced by chemosynthesis, and therefore it may prove to be extremely difficult to distinguish between a predominantly photosynthetic and a predominantly chemosynthetic basis for the near-vent food chains by measuring the local density of organic matter and associated bacteria, or by examining the spatial gradients in those densities. Even samples thought to be pure vent water[3] are not easy to interpret; the possibility of dilution or contamination by materials of non-vent origin will remain troublesome, because of the intense mixing around the vent[5].

Some of these uncertainties and paradoxes may soon be resolved by carbon-isotope data. The stable isotope ratio, ^{13}C to ^{12}C, depends on the nature of an organism's carbon sources, and is also a sensitive indicator of alternative metabolic pathways involved in photosynthesis and of heterotrophic fractionation. In the tissues of some of the vent fauna, the $^{13}C/^{12}C$ ratio differs significantly from values typical for other deep-sea animals[12–14], suggesting that some sort of unusual carbon sources or metabolic processes — perhaps chemosynthesis — may be important in near-vent areas. The finding that the ^{13}C ratios differ between vent bivalves and vestimentiferan worms in the direction of this anomaly (H. Craig, personal communication) is so far unexplained, the one group having smaller and the other larger values than expected. Also of great interest are data on the ratios of ^{14}C to ^{12}C in the tissues of near-vent animals, which ought to reflect their food resources, and which indicate[14] an apparent age of about 2,500 yr. Intra-vent chemosynthesis would incorporate the dissolved inorganic carbon of vent water, 80% or more of which apparently comes from magmatic sources (H. Craig, personal communication). If one assumes that magmatic CO_2 is very old and contains negligible amounts of ^{14}C, then in-vent chemosynthesis should lead to ^{14}C ages (>15,000 yr) very different from those obtained for the tissues of vent macrofauna ($\backsim$2,500 yr). So far, however, no actual measurements of ^{14}C in vent water are available, and other interpretations about its probable ^{14}C age have been proposed[14]. The apparent age of 2,500 yr of the circumvent fauna is clearly much older than concurrent sea-surface photosynthetic production (expected age $\backsim$0), and much younger than the available data for organic carbon incorporated into the top few millimetres of deep-ocean

The authors are professors at the Scripps Institution of Oceanography, La Jolla, California.

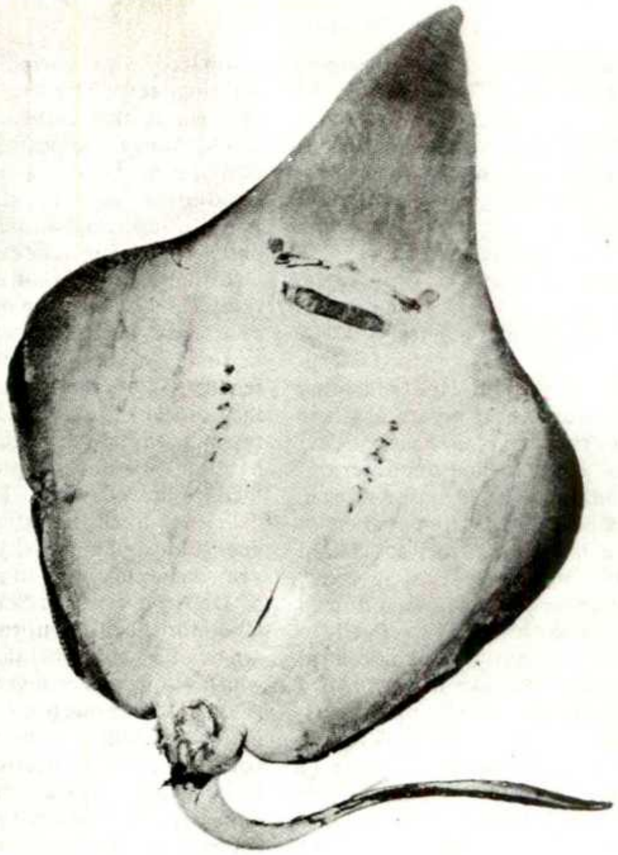
1. Lonsdale, P. *Deep-Sea Res.* **24**, 857 (1977).
2. Ballard, R.D. & Grassle, J.F. *Natn. Geogr.* **156**, 689 (1979).
3. Corliss, J.B. *et al. Science* **203**, 1073 (1979).
4. Galápagos Biology Expedition Participants *Oceanus* **22** no.2, 2 (1979).
5. Karl, D.M., Wirsen, C.O. & Jannash, H.W. *Science* **207**, 1345 (1980).
6. Killingley, J.S., Berger, W.H., MacDonald, K.C. & Newman, W.A. *Nature* **287**, 218 (1980).
7. Spiess, F.N. *et al. Science* **207**, 1421 (1980).
8. MacDonald, K.C., Becker, K., Spiess, F.N. & Ballard, R. *Earth planet. Sci. Lett.* (in the press).
9. Eittreim, S.L. & Ewing, M. in *Studies in Physical Oceanography Vol.2* (ed. Gordon, A.L.) 123 (Gordon & Breach, London, 1972).
10. Jacobs, M.B., Thorndike, E.M. & Ewing, M. *Mar. Geol.* **14**, 117 (1973).
11. Gardner, W.D. thesis, Massachusetts Inst. Technol. (1977).
12. Rau, G.H. & Hedges, J.I. *Science 203*, 648 (1979).
13. Rau, G.H. *Nature* (in the press).
14. Williams, P.M., Smith, K.L., Druffel, E.M. & Linick, T.W. *Nature* (in preparation).
15. Williams, P.M., Stenhouse, M.C., Druffel, E.M. & Koide, M. *Nature 276*, 698 (1978).

A new stingray from South Africa

from Alwyne Wheeler

ICHTHYOLOGISTS are accustomed to the regular description of previously unrecognized species of fishes, which if not a daily event at least happens so frequently as not to cause great comment. Previously undescribed genera are likewise not infrequently published, but higher categories are increasingly less common. The discovery of a new stingray, which is so different from all known rays as to require both a new family and a new suborder to accommodate its distinctive characters, is therefore a remarkable event.

A recent paper by P.C. Heemstra and M.M. Smith (*Ichthyological Bulletin of the J.L.B. Smith Institute of Ichthyology* **43,** 1; 1980) describes this most striking ray as *Hexatrygon bickelli* and discusses its differences from other batoid fishes. Surprisingly, this remarkable fish was not the result of some organized deep-sea fishing programme, but was found lying on the beach at Port Elizabeth. It was fresh but had suffered some loss of skin by sand abrasion on the beach, and the margins of its fins appeared desiccated in places. The way it was discovered leaves a tantalising question as to its normal habitat, but Heemstra and Smith suggest that it may live in moderately deep water of 400–1,000 m. This suggestion is supported by its general appearance (small eyes, thin black dorsal skin, flacid snout) and the chemistry of its liver-oil.



Ventral view of *Hexatrygon bickelli*

The classification of *Hexatrygon* presents some problems, the most obvious of which is that it has six pairs of gill openings, whereas all other batoid fishes (sawfishes, rays, skates, stingrays and electric rays) have only five paired gill openings. Another striking feature is that the spiracular openings behind the eyes are closed dorsally by a cartilage-supported flap to form an oblique slit. It seems that *Hexatrygon* may be able to shut its spiracles down externally, something that other batoids cannot do. Internal features are as strikingly different from its relatives, and the shape and structure of the snout are also unique. The snout is elongate, of uniform thickness (16–18 mm deep in the fish of 103 cm total length) and flaccid. Internally it is supported by lightly calcified rostral cartilages and filled with an acellular jelly, while the underside is richly supplied with well developed ampullae of Lorenzini. These ampullae are known to be electroreceptors and assist in finding buried food organisms in shallow-water rays. The combination of characters persuade Heemstra and Smith to erect a new suborder (Hexatrygonoidei) within the order Myliobatiformes to contain this fish. The order contains the other stingrays, eagle rays and mantas, most of which, like *Hexatrygon*, have one or more serrated dagger-like spines on the back of the tail.

The interest of this fish is not just that it represents a novel group within a well known order, but that all the other myliobatiformes are shallow-water bottom-living species or surface-living fish. *Hexatrygon* seems to be the first deep-water stingray known. Its elongate, jelly-filled snout richly supplied with electroreceptors is similar to the deep-water chimaeroids *Rhinochimaera* and *Harriota*, and there can be little doubt that it is an adaptation for finding food in the soft-bottom ooze of deep water.

One final point that Heemstra and Smith note is the remarkably small brain of *Hexatrygon* (it is about 3% of the cranial volume); in shallow-water stingrays it is about 80%. As they point out the coelacanth, *Latimeria chalumnae,* has a similarly small brain (as does the basking shark, *Cetorhinus maximus,* according to Scott *J. Fish Biol.* **16,** 665; 1980). This may also be an adaptation to deep-water life.

If nothing else the discovery of *Hexatrygon* off South Africa, as with the first coelacanth, reminds one of the tag — *ex Africa semper aliquid novi.*

Alwyne Wheeler is in the Department of Zoology, British Museum (Natural History), London.

---

sediment (apparent age $\sim$8,000 yr)[15]. For the advective contribution to vent nutrition, however, the appropriate comparison would be with the ^{14}C content of suspended particulate organic matter ordinarily present in near-bottom waters, which ought to be somewhere between these extremes, but again, appropriate measurements are not yet available.

In addition to the uncertainties associated with the two initial hypotheses for vent nutrition considered here, a third hypothesis has now entered the scene. Recent evidence (G.N. Somero and H. Felbeck, personal communication) indicates that significant bacterial chemosynthesis may be taking place *outside* the warm-water vents, by symbiotic microorganisms within the tissues of some of the macro-fauna. Hence, the basic problem of how the hydrothermal-vent fauna gets enough to eat remains an open and challenging question.

We are most grateful to the many colleagues who have shared their ideas with us, as well as those who have given us access to their unpublished data. □

0028-0836/81/040221-01$01.00

Untangling ataxia-telangiectasia

from Bryn A. Bridges and David G. Harnden

THE rare autosomal recessive disorder, ataxia-telangiectasia (AT), has attracted attention because it is associated with an unusual sensitivity to ionizing radiations and an increased susceptibility to cancer, particularly of the lymphoid tissues. Patients also have a variety of other problems, including a progressive neurological disorder, a substantial immunological deficiency, spontaneous chromosome instability, a defect of DNA repair, premature ageing, raised serum alpha fetoprotein(AFP) levels and other less regular features. How are all these diverse manifestations of the genetic defect interrelated? In an attempt to answer this question and to resolve inconsistencies in reported results, workers from a variety of different fields came together at a recent conference* at the University of Sussex.

Many laboratories have now confirmed the cellular radiosensitivity of AT cells to ionizing radiation and there was agreement that AT cells are also unusually sensitive to the chromosomally radiomimetic anti-tumour agent, bleomycin. Not all agreed that AT cells are sensitive to a range of other chemical mutagens, suggesting that increased sensitivity to any of these agents cannot be very marked although some of the disagreements may be due to differences between AT lines. Three different groups (Paterson, Chalk River; Inoue, Mishima; Jaspers, Rotterdam) all reported complementation between different AT lines for some radiosensitive function and agreed that AT3BI could be distinguished from AT4BI and AT5BI. Attempts to elucidate the molecular basis of the radiosensitivity failed to detect differences between AT cells and normals in the repair of single-strand or double-strand breaks in the DNA using presently available techniques. Two different types of abnormal molecular response were described, a failure to inhibit DNA replication after γ- or X-irradiation, and a defect in one or more repair pathways. A possibly analagous situation exists in *Escherichia coli* where mutations at the *rec A* gene block both inhibition of DNA replication after UV exposure and several pathways of DNA repair. The existence of a repair defect was demonstrated in different ways by several different groups. Cox (Harwell) showed a failure to repair potentially lethal damage in AT cells following X-irradiation. Indirect evidence of a repair defect comes from cytogenetic studies following G_0 irradiation of AT lymphocytes which reveal a large increase in fragments and chromatid aberrations as compared with normal cells (Taylor, Birmingham). This, together with the observation that caffeine can potentiate DNA damage to visible chromosome damage at least 24 hours after G_0 irradiation of AT cells (Natarajan, Leiden), suggests that there are long-lived unrepaired lesions. The exact nature of the DNA repair defect is still not clear and it is possible that more than one pathway may be affected. Paterson detected a reduction, in some AT cell lines,of γ ray-induced repair replication,and also a deficiency in their capacity to excise and repair specific γ ray-induced base damage, although there has been difficulty in reproducing some of these results elsewhere. Inoue reported a deficiency in cell-free extracts of all AT cell lines studied of an activity which modifies γ ray-damaged $3'$ sites to permit the action of DNA polymerase. Some, but not all, laboratories could confirm these results.

The other abnormal radiation response which was agreed by several groups is that AT cells show a much less marked decrease in DNA synthesis than do normal cells following γ-irradiation (Edwards, Birmingham; Lehmann, Sussex; Jaspers). Similarly, Scott (Manchester) finds that AT cells do not show the normal post-irradiation mitotic delay. It is not clear how these defects relate to the repair defect.

Unlike the other cancer-prone DNA repair deficiency syndrome, xeroderma pigmentosum, where there is a hypersensitivity of the cultured cells for mutagenesis by UV light, there is no evidence for any spontaneous or radiation-induced hypersensitivity for the induction of gene mutations in AT (Cox; Simons, Leiden). Furthermore, Arlett (Sussex) was unable to detect any radiation-induced gene mutation in AT cells. One of the most interesting features of AT patients is the high level of spontaneous chromosome aberrations in circulating lymphocytes and particularly the tendency of cytogenetically abnormal clones to proliferate in the absence of any haematological abnormality. The formation of these clones following breakage at specific points on chromosome 14 was emphasized and it was suggested (Hecht and McCaw, Tempe, Arizona) that aberrations involving the breakpoint at or near 14q32 may have particular neoplastic potential.

There is definite evidence for a profound immunological defect in AT patients. All have a defect of cell-mediated immunity and some have a deficiency of IgA and IgE. Waldmann (Bethesda) showed that the IgA deficiency is due to lack of synthesis rather than abnormal metabolic rates. There is diminished helper T-cell function, but even stimuli which do not require a helper T function fail to provoke an IgA response, suggesting a defect of B-cell maturation. Waldmann also presented evidence to suggest that the DNA sequences necessary for IgA production are present in AT patients and there was some speculation about the possibility that the cutting and splicing of DNA required to produce the specific immunoglobulin genes may be impaired. The *rec A* gene of *E. coli* also controls the cutting and splicing involved in genetic recombination. The relationship betwen the immune defect and the occurrence of lymphoid cancers was considered. Unlike the other immunodeficiency syndromes AT patients show a range of epithelial cell tumours in addition to the preponderance of lymphoid cancers (Spector, Minneapolis). The pattern is also different from that seen in irradiated populations, where there is a preponderance of leukaemias rather than lymphomas. No conclusion was reached, but it seems likely that the cancers in AT are not related in a simple way to either the radiosensitivity or the disordered immune system, but rather to some combination of the two.

The neurological defects in AT patients, particularly progressive cerebellar ataxia, could not at present be readily linked with the biochemical or immune defects. There was, however, evidence that differentiation may be imperfect in many tissues and organs especially gonads, liver, the immune system and the central nervous system and this may in some way be related to the repair defect. The uniformly raised AFP levels in these patients may represent some immaturity of the liver. The deficient thymus appears embryonic rather than atrophic. Kidson (Brisbane) suggested that failure of the nervous system to differentiate properly might in time lead to progressive loss of neurones.

There was a great deal of interest in the possibility of detecting individuals heterozygous for the AT gene. This stems at least in part from the observation that heterozygotes had an increased risk for largely the same spectrum of tumours as the homozygotes (Swift, North Carolina). If current estimates of the AT gene frequency are correct, heterozygotes would comprise one per cent or so of the general population and this would have profound implications. It could be, for example, that AT gene carriers may account for a substantial proportion of early cancers of the lymphoid system. Suggested methods of detecting heterozygotes were low-dose chromosomal radiosensitivity (Kidson) and cell survival under hypoxic conditions (Paterson). The differences between obligate heterozygotes and normals were small, but may be significant in some cases. However, overlap between these two groups means that those tests would not be suitable for the recognition of heterozygotes in the general population.

Bryn Bridges is Director of the MRC Cell Mutation Unit and a Professor at the University of Sussex. David Harnden is Professor of Cancer Studies at the University of Birmingham.

---

*The workshop was held from November 6th to 7th, 1980, and was organized jointly by the MRC Cell Mutation Unit and the Department of Cancer Studies, University of Birmingham. Sponsorship was provided by the MRC, the Cancer Research Campaign and the International Cancer Research Workshops programme (ICREW) of the UICC.

0028-0836/81/040222-02$01.00

The emphasis of recent research on AT has been on the sensitivity of these cells to various agents and their effect on DNA synthesis and repair and progress in these areas has been substantial. While these are undoubtedly important, more effort could in future be directed towards relating these areas to the defect of the immune system and the neurological disorder. Such a relationship, though surprising, must exist since the evidence that AT is due to a single autosomal recessive gene is good. ☐

The spruce invasion

from Peter D. Moore

ONE of the greatest weaknesses of palaeoecology is its necessary reliance upon circumstantial rather than experimental evidence. Yet it is this weakness which also opens the way to interminable speculation on the basis of the temporal correlation of events in the past. A very good example is the debate concerning the expansion of spruce (*Picea abies*) in Europe since the last glaciation.

Pollen analysis shows that, following the final glacial retreat, *Picea* immigrated into the eastern Alps very rapidly, accompanied by deciduous tree genera such as alder (*Alnus*), hazel (*Corylus*), oak (*Quercus*) and elm (*Ulmus*). It was preceded by birch (*Betula*), pine (*Pinus*) and larch (*Larix*). In Austria, for example, this deciduous forest type expanded into the south-east around 10,200 BP (Fritz *Carinthia* II **83**, 277; 1973) and by 9,500 BP it had expanded westwards in the Tyrol (Bortenschlager *Ber. naturw.-med. Ver. Innsbruck* **63**, 105; 1976). Although *Picea* was a member of the early post-glacial forests of Austria, even in the upper valley of the Inn, it is interesting to note that its immigration into Switzerland was considerably later; and it is here that scope for speculation begins.

Markgraf (*Nature* **288**, 249; 1970) collated the Swiss data concerning the movement of *Picea* into Switzerland and she found that its expansion was not very regular. Only in the south-east had it arrived by about 6,000 BP, but between that date and about 4,500 BP *Picea* pollen increased very markedly in a series of sites scattered all over the country, extending right to the western borders, north of Zurich. There remained some sites, however, even in the south of the country, in which the expansion was delayed until about 3,000 BP or even later. The expansion of spruce was often accompanied by a decrease in other tree species, including fir (*Abies*), pine and the deciduous forest genera. Also there was a simultaneous increase in herbaceous pollen, largely grasses, which suggested that the forests had been disturbed by human activity and that this provided spruce with an opportunity to establish itself where previously it had been excluded by competitive pressures. The combination of these pieces of circumstantial evidence led Markgraf to reject the earlier theories of a climatic change which provided the right conditions for *Picea* invasion into Switzerland and to propose that the tree followed in the wake of Neolithic agricultural progress.

One still needs to explain the delay in expansion between Austria and Switzerland. Here Markgraf suggested that the early invasion of *Abies*, *Larix* and *Pinus cembra* (the only central European member of the *Haploxylon* subgenus of *Pinus* and therefore distinguishable on the basis of its pollen) may have formed a dense canopy forest into which *Picea* could not immigrate. All three of these species were present in Austria, but either arrived later or were in lower frequency. It is claimed that such a forest may have a certain degree of inertia which renders it resistant to invasion by new species and also resilient in the face of climatic fluctuations (Smith *Proc. R. Soc.* **B161**, 331; 1965).

Spruce has also moved westwards in Fennoscandia in the latter half of the post-glacial, but here alternative explanations have been proposed. Moe (*Bot. Notiser* **123**, 61; 1970) collected the dates for the rise in *Picea* pollen in this area and showed that its expansion began in eastern Finland by 5,000 BP, reached the Baltic coast of Finland by 3,500 BP and subsequently spread through Sweden and into Norway. It has an unbroken distribution, reaching the west coast of Norway only in the central

New vistas in soil research

from Hamish Anderson

FRANCIS BACON'S assertion, that "there is nothing makes a man suspect much, more than to know little", gives an adequate summary of the present state of soil organic chemistry. Soils produce a [illegible] array of organic and [illegible], long [illegible] ultimate basis of [illegible] proof. There remain few separation techniques or analytical methods which have not been used in humic substance research, and the structural chemistry of soil extractives has uncertainly advanced, mainly by degradative techniques (see Schnitzer and Khan *Soil Organic Matter* Elsevier, 1978).

Proton magnetic resonance spectroscopy is an obvious tool but initial results were beset with line-broadening problems arising from chemical shift anisotropy, even in the rare cases where samples were sufficiently soluble for examination. The advance in NMR capability has now led to the threshold of the soil organic chemist's dream, the possibility of examining humic substances *in situ* (Barron *et al.* this issue of *Nature*, page 275). The Antipodean researchers have shown that the use of cross-polarization ^{13}C NMR techniques has allowed the [illegible] of useful spectra from whole [illegible]. Conventional ^{13}C spectra of amorphous solids usually give no useful structural clues, mainly due to large dipolar interactions. The ^{13}C cross-polarization technique leads to increased sensitivity by the elimination of heteronuclear dipolar broadening, this being achieved by transferring polarization from ^1H to ^{13}C (Mehring *High Resolution NMR Spectroscopy in Solids* Springer-Verlag, 1976). The transfer from the more abundant and more polarized ^1H spins leads to a possible quadrupling of the ^{13}C spins and, more importantly, the ^{13}C spin-lattice relaxation rate now occurs at that of the proton population. Since this is usually an order of magnitude faster than the conventional ^{13}C rate, the data accumulation frequency can be greatly increased. Even allowing for this 'speeding-up' process, it is sobering to realize that the number of free induction decays collected by Barron *et al.* during whole soil spectrometry ranged from 2×10^4 to 2×10^5. One major defect of the present system is that the chemical shift anisotropies of aromatic and carboxylic carbon lead to signal broadening and loss of resolution. However Barron *et al.* are confident that the combination of the present high-field cross-polarization technique with 'magic-angle' sample spinning will resolve this problem, as has been demonstrated with isolated humic substances (Hatcher *et al. Org. Geochem.* **2**, 87; 1980).

Humic substances research has long been dogged by the uncertainty of structural alterations occurring during extraction, and the proposed techniques offer answers to this problem. Indeed, the fact that modern NMR spectroscopy offers an ever-widening scope for examining a variety of nuclei instils hope that whole soils and their physical separates can be studied, resolving the interactions between humic substances and silicates, along with the involvement of nutrients such as nitrogen, sulphur and phosphorus.

Hamish Anderson is in the Department of Soil Organic Chemistry, Macaulay Institute for Soil Research, Craigiebuckler, Aberdeen.

0028-0836/81/040223-02$01.00

region. There are some similarities to the Swiss data in that the expansion is non-synchronous, but in this case there is a more distinct pattern of progressive westward movement.

Tallantire (*Norw. J. Bot.* **19**, 1; 1972) assessed the possible courses of the late post-glacial movements of spruce in Fennoscandia and he regarded the expansion as a step-wise development with phases of expansion, and plots them in this form on his maps. He quotes in full the dates upon which he bases this idea but, if the dates are considered in conjunction with their errors, it cannot be fully supported; it is possible to interpret the spread of spruce as a gradual process. Tallantire considers the expansion a consequence of climatic changes and bases this partly upon his assumption of phased movements, but also upon the inconclusive nature of the palynological evidence for coincident human activity. Further analysis of the climatic requirements of spruce (*J. Biogeogr.* **4**, 219; 1977) led Tallantire to regard the critical parameters as ample early-summer water supply and stable, low winter temperatures. In the absence of either of these factors, the young spruce (2–3 years) is selected against. He concludes "the conformable arrangement of the dates of spruce spread in southwestern Finland and its relationship to present winter climatic data is surely too marked to be due to a chance distribution of the investigated sites."

Since we are dealing with correlations, however, it is worth pointing out that correlation between pollen evidence for human activity and the spread of the spruce is becoming increasingly convincing as more detailed data accumulate, particularly from Finland. In the Kuusamo district of the north-east, *Picea* expanded at the time when tree pollen was falling and increasing Gramineae, *Calluna*, *Vaccinium* and *Empetrum* indicate ground flora changes (Hicks *Trans. Inst. Br. Geogr.* NS1, 362; 1976 and *Commentat. Biol.* **80**, 1; 1975). Although Hicks does not place such an interpretation upon the data, it is possible to support the idea of human agency in *Picea* spread.

Similarly in southern Finland, Tolonen (*Ann. Bot. Fenn.* **15**, 4; 1978 and **17**, 15; 1980) produces pollen stratigraphic curves in which the indications of human activity at the times of *Picea* expansion are strong. Many weed species, in addition to cereals, show short-duration peaks at precisely the time of the *Picea* rise at Lake Lammijärvi. At Lake Ahvenainen, *Picea* pollen rises in a series of episodes at each of which there are also cultural indicators. It is also noteworthy that considerable increases of charcoal are found in the latter site following the horizon at which *Picea* begins its spread.

Tallantire has claimed that the ecological conditions in Fennoscandia were essentially different from those in central and southern Europe in that *Abies, Fagus* and *Carpinus* were not found in the pre-*Picea* woodlands, whch consisted largely of *Betula, Pinus* and *Sorbus*. These forests would therefore have offered little competitive resistance to *Picea*, it may be supposed. The correlation between forest disturbance and spruce invasion certainly exists, however, and it may be that nutrient release from litter caused by disturbance and especially burning was influential in allowing spruce to invade. In the current management of Fennoscandian spruce forest, for example, Viro (in *Fire and* Ecosystems (eds T.T. Kozlowski & C.E. Ahlgren), Academic Press, New York, p.7; 1974) recommends the use of fire for rejuvenating old stands.

We are thus faced with the task of balancing alternative correlations, bearing in mind that a strong probability exists that both processes were operative. In the absence of any possibility of experimental confirmation of either hypothesis, the ground is well prepared for lengthy debate and speculation. ☐

Peter D. Moore is in the Department of Plant Sciences, University of London, King's College.

Inclusion compounds

from J.E.D. Davies

COMPOUNDS which show unusual properties, such as a preference for methanol rather than for ethanol, for the *para* rather than for the *ortho* or *meta* isomers of $C_6H_4Cl_2$, for the axial rather than for the equatorial conformer of $C_6H_{11}Cl$, for the *trans* rather than for the *gauche* rotamer of 1-chlorooctane, and even a preference of $C_6D_4(CD_3)_2$ rather than for $C_6H_4(CH_3)_2$, were the subject of a recent symposium* in Poland.

These interesting compounds are the inclusion compounds, which are two-component solid-state systems where the 'guest molecule' is accommodated in voids formed in the crystal structure of the 'host lattice'. These host–guest complexes (for which the term 'hogiplexes' was suggested [J.E.D. Davies, Lancaster]) can be classified according to the type of void available for the guest molecule. The guest species is accommodated in interlayer spaces in the intercalates; in the clathrate compounds the guest species is completely enclosed in a cavity; and the term inclusion compound is usually used where the guest species is accommodated in any other type of void, such as long channels (see figure).

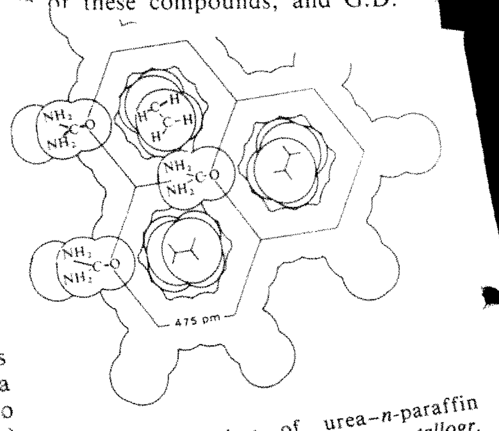
After a historical introduction [illegible] Kemula (Warsaw) and a general survey of the different types of host–guest complexes by J.E.D. Davies (Lancaster), the symposium heard about current work using a variety of host lattices.

The greatest number of papers was devoted to the Werner-type host lattices, M(X-py)$_4$ (NCS)$_2$, where M = Mn(II), Co(II), Ni(II), Fe(II) and X=3-Me, 4-Me, 4-Et, whose selectivity towards isomeric disubstituted benzenes has long been known (Schaeffer *et al. J. Am. Chem. Soc.* **79**, 5870; 1957) and recent selectivity studies were reported by S. Brzozowski (Warsaw), J.E.D. Davies (Lancaster) and A. Lewartowska (Warsaw). This selectivity can be put to good use, and D. Sybilska (Warsaw) showed how it is possible to separate isomers chromatographically on columns consisting of the Werner complexes. The chromatographic use of a wider range of inclusion compounds was reviewed by E. Smolkova-Keulemansova (Prague).

Undoubtedly the most unexpected selectivity shown by these host lattices is towards isotopic molecules. N.O. Smith (New York) described his careful work in this area and was able to show that when Ni(4-Mepy)$_4$(NCS)$_2$ is equilibrated with a mixture of the protonated and deuterated guest, there is a definite selectivity towards the deuterated species with benzene, xylene and naphthalene.

The related compounds, Ni(NCS)$_2$-(α-arylalkylamine)$_4$, also show remarkable selectivities and P. de Radzitzky (Brussels) described the work of the Labofina group on these compounds and illustrated a plant using these compounds to separate the dimethylnaphthalenes on an industrial scale.

Crystallography has been an important technique in the study of inclusion compounds, particularly since it was the pioneering crystallographic studies of H.M. Powell which first elucidated the [illegible] of these compounds, and G.D.



Cross-sectional view of urea–*n*-paraffin complex. (Fig.11 from Smith *Acta crystallogr.* **5**, 224; 1952. Reproduced by permission of *Acta crystallographica.*)

---

*First International Symposium on 'Clathrate Compounds and Molecular Inclusion Phenomena' organized by the Institute of Physical Chemistry of the Polish Academy of Sciences and held at Jachranka, Warsaw, 22–26 September 1980. The author acknowledges receipt of a British Council Travel Grant.

0028-0836/81/040224-02$01.00

The source of meteorites in time and space

from Robert Hutchison

VARIOUS false alarms have occurred in the past over alleged discoveries of fossil meteorites. For example, a squarish piece of iron weighing 785g from Tertiary lignite mined at Wolfsegg, Austria, was originally reported as meteoritic (*C.r. hebd. Séanc. Acad. Sci., Paris* **103**, 702; 1886; *Nature* **35**, 36; 1887). More recently, in 1930, a fist-sized piece of iron-nickel metal was said to have been recovered from a borehole at a depth of 1,525 feet in Eocene sediments; this 'Zapata County', Texas, iron has since been lost (Lovering *Nature* **183**, 1664; 1959). Until the discovery of the remains of a chondrite in Ordovician limestone in Sweden by Thorslund and Wickman (see this issue of *Nature*, p. 285), the only probable fossil meteorite material known to the writer was the occurrence of chondrule-like structures in Mesozoic sediment from the Urals (Yudin *Meteoritics* **6**, 99; 1971). The Ordovician chondrite has been reasonably (but not conclusively) identified as a member of the high-iron or H-group of the ordinary chondrites. Its fall to Earth some 460 Myr ago has considerable significance.

There are compelling arguments that most meteorites come from asteroids. But although it now seems probable that there are sufficient numbers of the Earth-crossing Apollo asteroids and the related Mars-crossing Amor asteroids to account for the observed flux of meteorites, a steady state can only be maintained if these groups are fed from some source beyond Mars (Wetherill *Geochim. cosmochim. Acta* **40**, 1297; 1976). The small Apollo objects have a mean lifetime of only a few tens of millions of years before they are either captured by a planet or ejected from the Solar System by Jupiter. About 30% of observed meteorite falls are H-group chondrites. Most of these meteorites have cosmic-ray exposure ages close to 4 Myr (see, for example, Zähringer *Geochim. cosmochim. Acta* **32**, 209; 1968), which is taken to mean that they were released as metre-sized objects from some larger body some 4 Myr ago. Release was presumably the result of a collision involving the parent body. No stony meteorite has an exposure age greater than 100 Myr. Thorslund and Wickman's discovery indicates that there was H-group chondrite material among Apollo asteroids some 460 Myr ago, and the possibility then arises that the meteorite flux has not changed significantly in composition with time.

If this is so there appears to be a problem. From spectral reflectance data Gaffey and McCord (*Proc. 8th lunar Sci. Conf.* **1**, 113; 1977) found that belt asteroids rarely have surfaces with the mineralogy of ordinary chondrites; such surfaces are common among Apollo and Amor objects. Wetherill concludes that to maintain a steady-state population of these bodies, belt asteroids and, dominantly, comets must have orbits which can evolve into Earth- or Mars-crossing orbits. However, belt asteroids appear to have surfaces which are almost exclusively not ordinary chondritic, and Anders (*Icarus* **24**, 363; 1975) has argued, on the basis of the ancient bombardment undergone by meteorites of different types by particles emitted by solar flares, that most stony meteorites cannot have come from the outer Solar System. The problem of the maintenance of the meteoritic flux was recently discussed by Napier and Clube (*Nature* **282**, 455; 1979), who proposed that a proportion of meteorites come from outside the Solar System, but the meteorites themselves have no major chemical or isotopic anomalies consistent with different regimes of nucleosynthesis; so we are left with a Solar System origin. The question of where in the Solar System that might be has, if anything, been aggravated by the new discovery.

Robert Hutchison is at the British Museum (Natural History), where he is responsible for the curation of the national collection of meteorites.

Andreetti (Parma) dealt with crystallographic studies of a whole range of inclusion compounds. It was therefore surprising to learn of the lack of crystallographic studies on Werner inclusion compounds and J. Lipkowski (Warsaw) reported his recent studies which aimed at trying to understand the factors governing host lattice selectivity. A knowledge of the crystal structure is also necessary to understand the energy-transfer processes between host and guest using UV and γ-ray radiation reported by the group from the Institute of Nuclear Chemistry, Rome. Another new term suggested at the meeting was 'excitrate' (A. Guarino, Rome) to denote a clathrate with either the host or guest, or both, in an excited state.

The Hofmann inclusion compound, $Ni(NH_3)_2Ni(CN)_4 . 2C_6H_6$, has been known since 1897, but its properties have only been recently studied extensively. The most prominent worker in this field, T. Iwamoto (Tokyo), described his work on modifying these host lattices by replacing the ammonia by bidentate bridging ligands, and by replacing the square planar $Ni(CN)_4{}^{2-}$ group by tetrahedral $M(CN)_4{}^{2-}$ groups. These modified host lattices show interesting properties such as the ability of $CdenHg(CN)_4$ to stabilize the cyclohexadienyl free radical, produced by the γ-ray irradiation of the benzene guest molecule. This free radical is stabilized up to the thermal decomposition temperature of the host lattice, 423K. Vibrational spectroscopic studies of these systems were reported by J.E.D. Davies (Lancaster) and stability studies were reported by A. Sopková (Košice).

It is true to say that the majority of inclusion compounds known have been discovered by accident, but D.D. MacNicol (Glasgow) gave a fascinating talk outlining his strategy for the synthesis of a variety of compouds which should be expected to act as host lattices. Equally fascinating was the talk by J.L. Atwood (Alabama) describing his work on 'liquid clathrate' systems. These are non-stoichiometric compounds which form in the liquid phase on the interaction of aromatic molecules and $M[Al_2R_6X]$ moieties. The mole ratio of aromatic to anion can vary enormously, the present record being $[N(C_6H_{13})_4][Al_2Et_6I].39.6$ *o*-xylene. A few of these liquid-phase systems have yielded crystalline products, for example $Cs[Al_2Me_6N_3]$ .2*p*-xylene, in which the *p*-xylene molecules sandwich the caesium ion.

A recent interesting application of inclusion compounds is the realization that a chiral host lattice might show enantiomer selectivity, and G. Tsoucaris (Chatenay-Malabry) gave examples of the enantiomeric selectivity of the tri-*o*-thymotide host lattice. Another recent development is the application of solid-state NMR spectroscopy to the study of inclusion compounds. J.A. Ripmeester and D.W. Davidson (Ottawa) illustrated the use of ^{129}Xe NMR in the study of the Structure I xenon clathrate deuterohydrate, where it was possible to observe well separated resonances from the atoms in the two types of cavity.

In conclusion, this was a well timed symposium, as it is apparent that the field of inclusion chemistry has advanced considerably in the 33 years since H.M. Powell first characterized this type of compound, and it is equally apparent that further advances remain to be made. Some of the papers presented at this meeting will be published in a special issue of the *Journal of Molecular Structure*, and it is hoped to hold another symposium in Parma in about two years time. ☐

J.E.D. Davies is in the Department of Chemistry, University of Lancaster, UK.

0028-0836/81/040225-01$01.00

Holographic bubble chambers

from David J. Miller

THE new particles have tantalisingly short lifetimes. Even the D^+, the longest-lived charmed particle, lasts for only about 10^{-12} seconds — time to travel 300 μm at the velocity of light, or a few millimetres with relativistic time dilation. But the D°, the Δ_{charm}, the F meson, the τ lepton and the 'beauty' mesons are predicted to travel only some tens of microns before decay. Nuclear emulsions have come back into popularity in the past five years because they gave the only chance of studying such short tracks, and now a new generation of tiny bubble chambers is developing which can resolve bubbles of less than 25 μm in diameter. Each technique has shortcomings — the emulsion requires many scanner-years of work to find the interesting events; a bubble chamber with conventional cameras has such a poor depth of field when the resolution drops below 25 μm that it becomes impossible to focus all of the beam particles into the sensitive volume.

At the Fermilab workshop* Le Coq (CERN) presented impressive pictures and videotapes that had been reconstructed from holograms made with a 6.5 cm diameter by 3.5 cm deep heavy liquid bubble chamber ('BIBC', the Berne Infinitessimal Bubble Chamber). The holographic technique gave clearly resolved images of bubbles only 8 μm in diameter over the whole depth of the chamber. This work has inspired a number of other developments. Pless (MIT) presented scaling rules for the change of resolution when Le Coq's 'straight-through' holographic layout is modified to medium-sized bubble chambers with $\sim$ 1 m diameter. His collaboration is now presenting a proposal to Fermilab for a study of the τ neutrino in such a chamber, using a neutrino beam from a beam dump. A proposal by a European collaboration for a small holographic hydrogen bubble chamber was discussed by Fisher (Rutherford). This would be a continuation of the work being done with conventional optics in proton and pion beams at CERN. He warned that, although the desired resolution can almost certainly be achieved over a useful volume, and although there are techniques available for detecting, measuring and identifying the secondary particles outside the chamber, one serious problem remains for the exploitation of holographic bubble chambers — a very similar problem to that faced by nuclear emulsions. The new particles are produced very rarely — in less than one event in a thousand. To make an emulsion, or a hologram, really useful, an electronic trigger must be devised which will identify the subsample of events which might contain the short-lived particle. The tracks from these events must then be extrapolated back into the emulsion or the chamber so that the scanner knows where to look. With a bubble chamber it should be possible, in principle, to pulse a laser and make the hologram only when a good event has been detected, but Fisher warned that no technique yet exists which can separate the rare particles cleanly from the background within the few microseconds that it takes for bubbles to grow to the desired size.

Some people think that an alternative approach might bring good results. Instead of looking for a 'needle in a haystack', using small chambers with fine resolution in proton, pion or photon beams, where the backgrounds are enormous, it is perhaps more sensible to work in a neutrino beam where charmed particles occur in 10 per cent of events. To have a reasonable number of neutrino interactions a large bubble chamber must be used, such as BEBC at CERN or the 15 foot chamber at Fermilab. At present these chambers grow their bubbles to around 600 μm in diameter in order to give a clear conventional image from more than 10 m^3 of sensitive volume. A few D^+ tracks have been resolved — the very furthest tail of the length distribution — but many more could be seen if 50 μm resolution were possible. Very little work has been done on the possibilities of using holography in the large bubble chambers but no one has demonstrated that it cannot be done. A number of methods were suggested: bright field 'Scotchlite' illumination, though work at Imperial College has shown that speckle may be a problem; dark field with Scotchlite, but can sufficiently powerful lasers be found; retrodirecting mirrors, if a suitable reference beam can be set up. Peterson (Hawaii) drew attention to the relevance of these ideas to proposals his collaboration is making for a new chamber at Fermilab. Its primary purpose is to allow secondary particles from neutrino interactions to leave the bubble chamber and be identified, but such a new device could also be designed to incorporate whatever is found to be the most favourable holographic layout.

We were reminded by Lederman, the Fermilab director, that finer resolution means smaller equipment which costs less to build. Fine electronic counters would enable experiments of all kinds to be scaled-down proportionally to the resolution, giving more than proportional reductions in magnet, material and building costs. Such counters were outside the scope of the workshop, but more than one group is convinced that the solution to the 'needle in a haystack' problem will depend on an emulsion or small holographic bubble chamber being followed by electronic counters with a resolution of 20 μm which will both improve the trigger and allow very precise predictions to be passed on to the scanners.

The workshop ended with a theorist's checklist of the effects that could be studied at high resolution. Bjorken (Fermilab) stressed that once the new particles can be collected in reasonable numbers they should show as complex and instructive a phenomenology as did the strange particles. □

D.J. Miller is a lecturer in the Department of Physics and Astronomy at University College London, at present on sabbatical leave in Fermilab.

---

*A workshop on holographic bubble and streamer chamber techniques was organized by L. Voyvodic at Fermilab, Batavia, Illinois on 11–12 November 1980.

100 years ago

Mr E. T. Sachs sends us some interesting notes from Batavia: — "Within the past month I have been so lucky as to make what I hope is a very interesting if not remarkable discovery. On the Island of Biliton, 200 miles from here, I found a fresh water fish which produces its young *living from its mouth.* I am quite prepared for the cry of incredulity that will be raised; but I conducted my observations with living fish and closed doors, and what I assert is undeniable: the eggs are hatched in the lower portion of the head of the fish, and are projected out at the mouth and from nowhere else. I have secured several specimens, which I shall send to Dr. Günther, who will of course at once set the matter at rest."

A "Natural" Experiment in Polarised Light

Break off a plate of ice and hold it between the sky and a pool of water. Its reflected image will show the beautiful colours due to polarised light. The incident rays should come from a part of the sky about 90° from the sun, and reflection should take place at the polarising angle for water, and the plate will probably require adjusting to bring out the maximum effect. Water, vaporous, solid, and liquid, thus furnishes us with polariser, crystal, and analyser. I do not remember to have read any account of this very simple experiment, for which Nature provides all the materials.

Chas. T. Whitmell

9, Beech Grove, Harrogate, January 10

From *Nature* **33,** 20 January, *268, 275,* 1881.

0028-0836/81/040226-01$01.00

Contaminated cell lines

from David Dickson, Washington

YET another corruption of the scientific literature has been uncovered with the discovery that several cell lines widely quoted in the scientific literature as originating from patients with Hodgkin's disease have now been shown not to be human cells at all, but to come from a strain of owl monkey (see the accompanying article, pages 228–230).

The cells were grown in culture by Dr John C. Long, who resigned a year ago as assistant pathologist at the Massachusetts General Hospital and associate professor of pathology at Harvard Medical School, after admitting to having faked data in a research paper published in the *Journal of the National Cancer Institute* in 1979.

On resigning Dr Long admitted having misled colleagues about carrying out experiments to produce revised data on the molecular weights of immune complexes. The alterations were included in the paper after it had been rejected by the referees of another journal.

Dr Long said on the telephone last week that during his time at the Massachusetts General Hospital, he had no reason to suspect that his cell lines were contaminated and welcomed the work of Harris *et al.* However, when he resigned, he did not reveal that six months earlier he had received data from the National Cancer Institute (NCI) casting doubt on the identity of the cells.

The first successful long-term cultivation of cells from patients with Hodgkin's disease was claimed by Dr Long and colleagues from the hospital and the medical school in *Proceedings of the National Academy of Sciences (PNAS)* in 1973. Other groups have encountered considerable difficulty in establishing permanent lines from diseased spleen cells.

The procedures used to grow the original cells through over 200 passages were described in a 1977 paper in *Journal of Experimental Medicine* as well as *PNAS*. (*Science Citation Index* lists 30 citations of this paper in 1979 and 1980, which its compilers say means the paper is "highly cited".) In total four papers have appeared in *PNAS* — three naming Dr Long as the first author. Experiments using these cells have also been described in other publications (see page 228).

All the papers referring directly to the now-challenged cell cultures will have to be reassessed to see if any data of scientific value can be salvaged, since there is now evidence that the cell lines were contaminated from an early stage.

Dr Long was awarded a three-year grant for more than $200,000 in 1976 to continue his study of the Hodgkin's disease tissue cultures. This was renewed in 1979 — based on an application said to include the data which he later admitted forging — and he received $96,000 in that year, as the first part of a three-year grant of almost $500,000.

This grant was terminated by the NCI at the request of the hospital, after Dr Long admitted falsifying data in the 1979 paper in *Journal of the National Cancer Institute*. According to the hospital, Dr Long claimed at the time that "competition for federal research grants" was a major factor in causing him to lie about experiments he had not carried out.

However, he did not admit to any other impropriety and the hospital put out a statement saying that the data he had admitted falsifying was "obscure, and of no known significance. . . no known harm has resulted even in the field of science".

The hospital also said that further suspicions about his research were being followed up. There was already evidence of contamination of the cell lines — early in 1979 one of Dr Long's research associates had grown suspicious about the true nature of the cells, and sent samples to Dr Stephen O'Brien at the NCI Laboratory of Viral Carcinogenesis for genotypic evaluation.

The NCI scientists told Dr Long that three of the cell lines carried the enzymatic signature of the same individual. And they also pointed out an unusual characteristic they could not explain — a peculiarly fast mobility in a key enzyme system.

Dr Long did not follow up this comment with the NCI. And he told his research workers that the NCI check had shown that the cells were genuine — indeed that in one case the signature of the cultured cells corresponded directly with that of the red blood cells of the patient from whom the cells were said to have been taken (another claim subsequently shown to be false).

Nothing further happened at the time. But after Dr Long's resignation, Dr Robert T. McCluskey, head of the hospital's pathology department, considered further investigations necessary. Karyotypes and cell samples were sent to the cell culture laboratory at the University of California's School of Public Health, where chromosomal analysis revealed that the cells were not human. Several other laboratories were subsequently sent samples to complete the identification. And the karyotype was recognized as that of the owl monkey by Dr Bharati Hukku of the Child Research Center of Michigan, Detroit.

The cells were then confirmed by the New England Regional Primate Center as coming from a Northern Columbian brown foot owl monkey. Indirect evidence that this was the likely source of contamination was the fact that this cell line had been used for virus research in the early 1970s in the laboratory in which Dr Long was then working.

The fourth suspected cell line, started later than the other three, has turned out to be human. But there is no evidence of a link with Hodgkin's disease tumours.

The full impact of Dr Long's publications remains unclear. His results have been quoted as supporting a particular view of the pathogenesis of Hodgkin's disease which may now be questioned, in particular suggesting a macrophage origin of the peculiarly large Reed–Sternberg cells found in Hodgkin's patients.

Other research workers have been careful not to put too much weight on Dr Long's results, remaining sceptical in the light of their own failure to culture the same cells. And Dr Henry Kaplan of Stanford University feels that the main impact may have been to reduce the incentive for others to tackle the problem of growing the cells *in vitro*, since Dr Long had already claimed success.

Investigators for the Department of Health and Human Services in Boston must now decide whether there is any evidence that should be passed to the Department of Justice to charge Dr Long with fraud against the federal government. Under the US Code, consciously providing false information when applying for federal assistance is a criminal offence.

The journals which published Dr Long's papers now have to decide what action to take. A representative of the National Academy of Sciences has said that the situation was "unique" for *PNAS*, which has never previously had to disown any paper it has published.

Papers submitted to *PNAS* are not formally reviewed if a member of the academy is a co-author. In Dr Long's case, each of the four papers that he published was written with academy member Dr Paul Zamecnik, who recently retired as professor of oncological medicine and director of the J.C. Warren Laboratories at Harvard Medical School, where Dr Long was a research student.

The National Institutes of Health are already discussing how to tighten up on evaluating the scientific content of research funded by federal grants. An audit of Dr Long's grant revealed no fiscal irregularities in the way that the money had been used, but did not go into the legitimacy of the research.

Ironically, the recently created President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research is now considering another Boston case, in which a research worker at the Boston University Medical School was shown to have faked data in a cancer survey.

Research administrators admit that cases of fraud — usually relatively minor — are not unusual, although few reach the public eye. "Whatever you see is probably only a small fraction of what actually goes on," according to Dr Donald Kennedy, president of Stanford University.

0028-0836/81/040227-01$01.00

Contamination of Hodgkin's disease cell cultures

Nancy L. Harris*, David L. Gang*, Steven C. Quay*, Sibrand Poppema*, Paul C. Zamecnik†, Walter A. Nelson-Rees‡ & Stephen J. O'Brien§

* Department of Pathology, Harvard Medical School, Immunopathology Unit, Massachusetts General Hospital, Boston, Massachusetts 02114
† Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545
‡ Cell Culture Laboratory, University of California School of Public Health, Naval Biosciences Laboratory, Oakland, California 94625
§ Laboratory of Viral Carcinogenesis, National Cancer Institute, Bethesda, Maryland 20014

Several laboratories have recently reported the establishment and characterization of long-term cell lines thought to be related to the neoplastic cell of Hodgkin's disease. Here, Harris et al. discuss evidence that some of these lines are, in fact, not related to Hodgkin's disease but are non-human contaminants.

IN an effort to determine the origin of the malignant cell of Hodgkin's disease, many investigators have attempted to establish permanent cell lines from tissues involved with the disease. Early attempts to culture tissues involved with Hodgkin's disease yielded either mixed cultures which could not be maintained *in vitro*, or Epstein-Barr virus-infected lymphoblastoid cell lines (see ref. 1 for a review). Recently, several laboratories have reported the establishment and characterization of long-term cell lines thought to be related to the neoplastic cell of Hodgkin's disease[2–7]. Our studies indicate that several of these cell lines are not related to Hodgkin's disease, but are non-human contaminants.

Long *et al.* reported the establishment of four permanent monolayer cell lines, designated FQ, RB, SpR, and RY, from tumour nodules in spleens of patients with Hodgkin's disease[7]. These four cell lines were considered to be neoplastic by virtue of their capacity for indefinite replication *in vitro*, aneuploidy and the production of invasive tumours on subcutaneous injection into nude mice[8]. Although the cells had no surface markers demonstrable by standard immunological techniques, several features, namely the production of lysozyme and esterases, the ability to phagocytose iron particles[9] and the ultrastructural morphology of the cells in culture[10], were felt to be consistent with a macrophage origin of the cultured cells. The establishment of these permanent cell lines, which seemed to be neoplastic and had features of macrophages, supported the

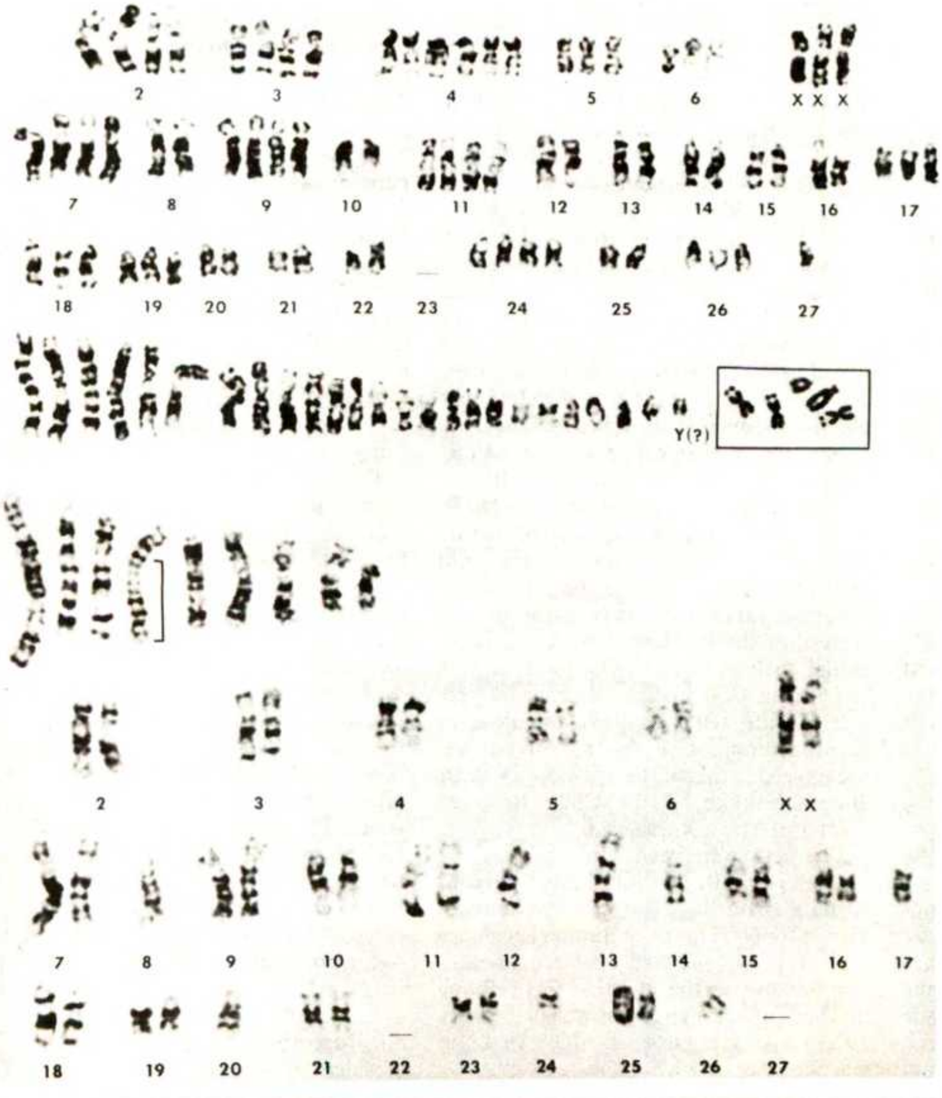


Fig. 1 Three marker chromosomes, conventional staining (insert). Trypsin-Giemsa banded metaphase chromosomes of an FQ (top four rows) and an OMK-210 cell (bottom four). The FQ culture exhibited polysomy for many chromosomes including, in some instances, duplicate marker chromosomes. OMK-210 cells (kindly provided by Dr M. Daniel) were near-diploid. Identical marker chromosomes in both cells are indicated by arrows. Most noticeable are isochromosomes for 13q. A small, altered chromosome which resembles the Y chromosome was noted among the markers in the FQ cells [designated Y(?)]. As the karyotype of OMK-210 is female, and the presence of multiple identical markers in the two cell lines indicates a common origin, it is unlikely that the small FQ chromosome is a true Y chromosome. The FQ karyotype was kindly constructed by Dr N. S. F. Ma; its pattern was followed for the OMK-210 cell.

0028-0836/81/040228—03$01.00

concept that the malignant-appearing cells observed by pathologists in Hodgkin's disease tumours are neoplastic macrophages[1,9,10].

Further studies of these cell lines appeared to offer a new insight into the pathogenesis of Hodgkin's disease. It was reported that two of these cell lines (FQ and RB) specifically bound immune complexes from the sera of patients with Hodgkin's disease[11,12]. As the cells lacked receptors for immunoglobulins and complement and did not bind immune complexes from sera of patients with other disorders, it was concluded that the antibody portion of the immune complex was reacting with an antigen on the cultured cells. This observation not only provided the first direct evidence that the cells cultured *in vitro* were related to the human disorder, but also suggested that some patients with Hodgkin's disease were reacting immunologically to a tumour-associated antigen.

A recurring problem associated with the use of long-term, established cell lines in the study of human tumours is the risk of contamination and overgrowth of cultures by unrelated cell lines[13]. Previous publications have shown that the techniques of trypsin-Giemsa-banded karyotyping[14] and isoenzyme (allozyme) analysis[15,16] can detect both intra- and interspecies contamination of cell lines. In addition to these methods of genetic characterization of cell lines, the technique of subcutaneous transplantation into nude (athymic) mice can also be used to investigate the origin of a cell line. When noncultured tumour cells or established cell lines are injected into nude mice, the tumours which form may be more differentiated than either the original tumour or the *in vitro* cultured cells, thus permitting determination of the histogenesis of an apparently undifferentiated cell[17].

Using these three techniques, we have found that three of the four cell lines (FQ, RB and SpR) reported by Long *et al.*[7] are identical to each other and are of owl monkey rather than human origin. Furthermore, they have morphological features of epithelial cells rather than macrophages. The fourth cell line (RY) is of human origin, but its relationship to Hodgkin's disease has not been established.

Cultures of FQ, RB, SpR and RY cells maintained by Dr J. Long, (Immunopathology Laboratory, Massachusetts General Hospital) were studied, as were the earliest available cultures from the laboratory of Dr P. Zamecnik (spanning passages 14-500 of FQ, 300-380 of RB, 21-220 of SPR, and 5-130 of RY). In FQ, RB and SpR cells the chromosome numbers ranged from 75 to 100, with a modal number between 83 and 100 chromosomes. Conventional chromosome preparations of FQ, RB and SpR metaphases revealed three to four copies of a medium-size submetacentric chromosome with a pronounced near-terminal secondary constriction; a chromosome peculiar to normal owl monkey cells (Fig. 1, insert). Trypsin-Giemsa-banded karyotopes of FQ, RB, and SpR exhibited non-human chromosomes, singly or in multiples, and were recognized as polyploid derivatives of cells of the owl monkey, *Aotus trivirgatus* ($2n = 54$)[18,19]. Cells at both early and late passages in culture contained many normal chromosomes of the owl monkey karyotype, as well as highly altered, marker chromosomes. Most notable among the latter were one or two isochromosomes involving the long arm of chromosome number 13 (Fig. 1). A search for an owl monkey cell line that was well adapted to *in vitro* culture led to the study of OMK-210 cells, an established cell line derived from owl monkey kidney[20]. Comparison of FQ, RB and SpR with OMK-210 cells showed that identical marker chromosomes, as well as normal chromosomes, existed in all lines.

Despite the presence of identical marker chromosomes, further chromosomal alterations had clearly taken place following the separation of FQ, RB and SpR from OMK-210 cells, leading to differences in their individual karyotypes. For example, although OMK-210 contained near-diploid numbers of chromosomes, the FQ, RB and SpR cells were all near-tetraploid. Such chromosomal alterations are common occurrences in cultured cell lines[21].

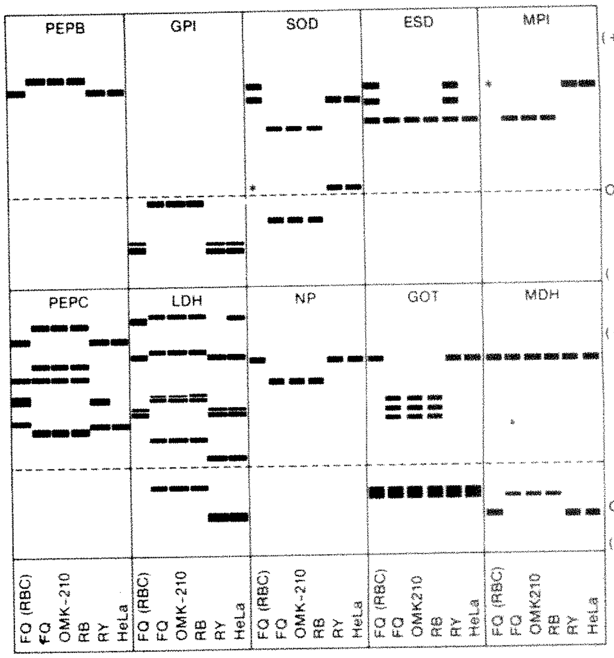


Fig. 2 Composite electropherograms of 10 gene–enzyme systems diagnostic for interspecies cell identification (refs 15, 16). The columns represent: 1, FQ patient's red blood cells; 2, FQ cell line; 3, OMK-210 cells; 4, RB cells; 5, RY cells; 6, HeLa cells. The 10 diagnostic enzymes were: PEPB, peptidase-B (EC 3.4.11.); GPI, glucose phosphate isomerase (EC 5.3.1.9); SOD, superoxide dismutase (EC 1.15.1.1); ESD, esterase-D (EC 3.1.1.1); MPI, mannose phosphate isomerase (EC 5.3.1.8); PEPC, peptidase-C (EC 3.4.11); LDH, lactate dehydrogenase (EC 1.1.1.27); NP, purine nucleoside phosphorylase (EC 2.4.2.1); GOT, glutamate oxaloacetate transaminase (EC 2.6.1.1); and MDH, malate dehydrogenase (EC 1.1.1.37). * Indicates that no activity was observed in that extract for the enzyme. All the systems except ESD vary between man (HeLa) and owl monkey (OMK-210). In man ESD is polymorphic for two alleles *ESD*-1 and *ESD*-2. The RY line is heterozygous for *ESD*-1-2. The FQ red cell extract shows that the FQ patient is heterozygous for *ESD*-1-2 while the FQ cell has an ESD with a mobility of *ESD*-1, which is identical to that of OMK-210. The OMK-210 isozymes were each identical to those observed in extracts of OMT-1, an owl monkey lymphoid cell line derived by Drs Louis Qualtiere and Gary Pearson (data not shown).

The identity of the FQ, RB and SpR cell lines with OMK-210 cells was further established by determining the electrophoretic mobilities of a panel of 10 isoenzymes selected to detect interspecies contamination (Fig. 2). The FQ, RB and SpR cell lines, at the passages examined, were identical to each other and to OMK-210 cells at each locus, and distinct from the human cell line, HeLa, and the red blood cells of the patient from whose spleen the FQ cell line purportedly arose.

Together, the chromosome and isoenzyme analyses indicated that the three supposed Hodgkin's disease cell lines had been cross-contaminated by OMK-210 cells. OMK-210 cells had, in fact, been carried in the laboratory of Drs Long and Zamecnik at the time when the FQ, RB and SpR cell lines were initiated. In the original description of the cell lines, it was reported that the Hodgkin's disease monolayer cultures underwent a morphological alteration *in vitro* after 3–10 months' propagation in culture, with the appearance of polygonal cells that replicate in a mosaic pattern[8]. This appearance of cells with different morphology from that first noted in the primary cultures or early passages is a characteristic occurrence in the course of contamination of cell lines[14].

Subcutaneous tumours formed by the FQ, RB and SpR cells in nude mice contained acinar structures with microvilli and intercellular junctions (Fig. 3*a*). These features indicated that the cultured cells were epithelial, and not of monocyte-macrophage or lymphoid origin[22]. The ultrastructural features of the nude mouse tumours contrasted with those of the cells when examined in monolayer culture[10], and indicated that additional differentiation had occurred in the nude mouse. Surface microvilli were present on the monolayer cells, but acinus formation and intercellular junctions were not seen *in vitro*[10]. Although a previous report on the transplantation of the FQ, RB and SpR

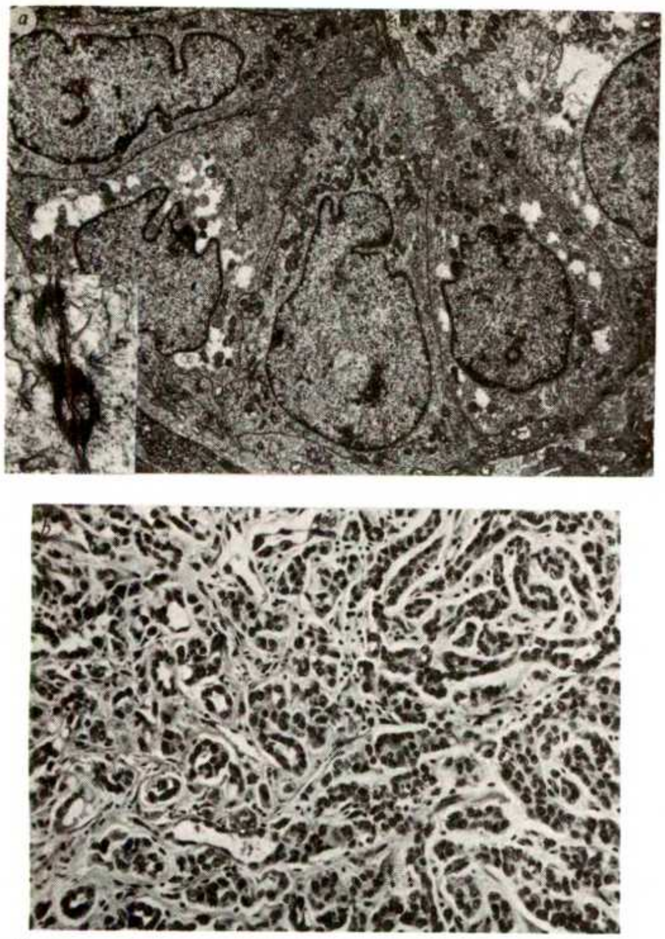


Fig. 3 *a*, Electron micrograph of tumour formed 3 weeks after subcutaneous injection of 5×10^6 FQ cells into a 6-week-old nude mouse (CRL:*nu*/*nu* (CD-1) BR, Charles River), fixed in Karnovsky-II and stained by the OCUB method[9]. ×2,750. Insert: intercellular junctions. ×27,500. *b*, Paraffin-embedded section of subcutaneous tumour formed in nude mouse by FQ cells (from J. C. Long, as described in ref. 7) showing gland and tubule formation (haematoxylin-eosin, ×185).

cell lines into nude mice stated that the morphology was that of 'reticulum cell sarcoma'[8], review of the original slides from that study revealed that tubular and acinar structures indicative of epithelial differentiation were present in the tumours formed by these three cell lines (Fig. 3*b*). This observation indicates that contamination of these cell lines had already occurred by then.

The origin of the fourth cell line, RY, could not be established with certainty. The cells had approximately 60 chromosomes. Conventional chromosome preparations did not contain the characteristic owl monkey chromosome with the near-terminal secondary constriction which was seen in the other three cell lines. Trypsin-Giemsa banding revealed an abnormal human karyotype, which was distinct from HeLa or any other established cell line known to us. The isoenzyme genetic signature likewise indicated that the cell line was human (Fig. 2); analysis of a different panel of isoenzymes, selected for their polymorphism within human populations[15,16], confirmed that RY differed from HeLa or any other known cell line (not shown). Tumours formed in the nude mouse by the RY cells, in contrast to the other cell lines, did not have epithelial features; the histogenesis of this cell line could not be determined from the morphology of the nude mouse tumours. As previously described[7], the RY cell line did not have surface receptors characteristic of macrophages. We were unable to confirm the initial report of lysozyme production by this cell line (S. Quay and S. Poppema, unpublished data). A review of the donor patient's hospital record revealed that the spleen from which the RY cell line originated contained no grossly or microscopically identifiable tumour. As this cell line fails to meet a basic criterion for a tumour cell line—namely, that it be cultured from a specimen known to contain tumour cells[23]—its relationship to the malignant cell of Hodgkin's disease remains in doubt.

Although a few other laboratories have reported the establishment of permanent cell lines from patients with Hodgkin's disease, no conclusions regarding the nature of the malignant cell have been drawn from them[2-4]. Although there are indications from other sources that the malignant cell of Hodgkin's disease is of macrophage origin, the evidence is not conclusive (for a review, see ref. 24). As the only permanently established Hodgkin's disease cell lines with macrophage characteristics, the cell lines reported by Long *et al.* were thought to provide substantial evidence for the macrophage origin of the Reed–Sternberg cell. Our findings, therefore, remove some of the support for the theory that the malignant cell of Hodgkin's disease is a macrophage. Furthermore, results obtained with these cell lines cast doubt on the specificity of the properties which were considered to indicate that these cultured cells were macrophages. Recently, the RY cell line was found to produce endogenous pyrogen, an observation which it was thought might explain the tendency of some patients with Hodgkin's disease to develop fever[25]. The significance of this observation must be re-evaluated in the light of our findings.

Our results also indicate that the establishment of permanent cell lines from tissues involved with Hodgkin's disease may be more difficult than had been previously suggested[6-9]. Several major laboratories have experienced repeated failure in their attempts to establish permanent cell lines from tissues involved with Hodgkin's disease[24,26,27]. The inability of these other workers to establish permanent lines *in vitro*, despite numerous attempts, may mean that appropriate methods for culture of these cells have not yet been found. On the other hand, it is possible that their very resistance to *in vitro* propagation has some biological significance which has not yet been recognized.

We specially thank Dr Bharati Hukku of the Child Research Center of Michigan, Detroit, who first noted the similarity between certain chromosomes in FQ and *Aotus trivirgatus*. Dr Ward D. Peterson Jr provided additional immunological and serological proof of contamination. Messrs R. R. Flandermeyer and D. W. Daniels and Dr L. Atkins assisted with karyology as did Dr N. S. F. Ma, whose classification of *Aotus* karyotypes was followed. We also thank Drs Robert T. McCluskey, Sigmund A. Weitzman, C. Alan Brown and Clive L. Hall for their support.

Received 10 November; accepted 17 December 1980.

1. Kaplan, H. S. *Hodgkin's Disease* (Harvard University Press, Cambridge, 1980).
2. Boecker, W. R., Hossfeld, D. K., Gallmeier, W. M. & Schmidt, C. G. *Nature* **258,** 235–236 (1975).
3. Roberts, A. N., Smith, K. L., Dowell, B. L. & Hubbard, A. K. *Cancer Res.* **38,** 3033–3043 (1978).
4. Schaadt, M., Fonatsch, C., Kirchner, H. & Diehl, V. *Blut* **38,** 185–190 (1979).
5. Long, J. C., Aisenberg, A. C., Zamecnik, M. V. & Zamecnik, P. C. *Proc. natn. Acad. Sci. U.S.A.* **70,** 1540–1544 (1973).
6. Long, J. C., Aisenberg, A. C. & Zamecnik, P. C. *Proc. natn. Acad. Sci. U.S.A.* **71,** 2285–2289 (1974); **71,** 2605–2609 (1974); *J. natn. Cancer Inst.* **58,** 223–227 (1977).
7. Long, J. C., Zamecnik, P. C., Aisenberg, A. C. & Atkins, L. *J. exp. Med.* **145,** 1484–1500 (1977).
8. Zamecnik, P. C. & Long, J. C. *Proc. natn. Acad. Sci. U.S.A.* **74,** 754–758 (1977).
9. Long, J. C. *Clin. Haemat.* **8,** 531–565 (1979).
10. Gang, D. L. Long, J. C., Zamecnik, P. C., Chi, S.-Y. & Dvorak, A. M. *Cancer* **44,** 543–557 (1979).
11. Long, J. C. *et al. New Engl. J. Med.* **297,** 295–299 (1977).
12. Long, J. C., Dvorak, A. M., Quay, S. C., Stamatos, C. & Chi, S.-Y. *J. natn. Cancer Inst.* **62,** 787–797 (1979).
13. Stulberg, C. S. in *Contamination in Tissue Culture* (ed. Fogh, J.) 1–27 (Academic, New York, 1973).
14. Nelson-Rees, W. & Flandermeyer, R. P. *Science* **191,** 96–97 (1976); **195,** 1343–1344 (1977).
15. O'Brien, S. J., Kleiner, G., Olson, R. & Shannon, J. E. *Science* **195,** 1345–1348 (1977).
16. O'Brien, S. J., Shannon, J. E. & Gail, M. H. *In Vitro* **16,** 119–135 (1980).
17. Sharkey, F. E. *et al.* in *The Nude Mouse in Experimental and Clinical Research* (eds Fogh, J. & Giovanella, B. C.) 187–214 (Academic, New York, 1979).
18. Miller, C. K. *et al. Cytogenet. Cell Genet.* **19,** 215–226 (1977).
19. Ma, N. S. F., Jones, T. C., Miller, A. C. Morgan, L. M. & Adams, E. A. *Lab. Animal Sci.* **26,** 1022–1036 (1976).
20. Melendez, L. V. *et al. Lab. Animal Care* **19,** 372–377 (1969).
21. Nelson-Rees, W. A. in *Origin and Natural History of Cell Lines* (ed. Barigozzi, C.) 25–79 (Liss, New York, 1978).
22. Mackay, B. & Osborne, B. M. in *Pathology Annual* 359–405 (Raven, New York, 1978).
23. Fogh, J., Goodenou, M., Loveless, J. & Fogh, H. in *In Vitro Methods in Cell-Mediated and Tumour Immunity* (eds Bloom, B. R. & David, J. R.) 677 (Academic, New York, 1976).
24. Kaplan, H. S. *Cancer* **45,** 2439–2474 (1980).
25. Bodel, P. Ralph, P., Wenc, K. & Long, J. C. *J. clin. Invest.* **65,** 514–518 (1980).
26. Kaplan, H. S. & Gartner, S. *Int. J. Cancer* **19,** 511–525 (1977).
27. Sundstrom, C. & Nilsson, K. *Acta path. microbiol. scand.* **86,** 173–184 (1978).

ARTICLES

Multiple asperity model for earthquake prediction

M. Wyss*, A. C. Johnston† & F. W. Klein‡

* Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309
† Tennessee Earthquake Information Center, Memphis State University, Memphis, Tennessee 38152
‡ Hawaiian Volcano Observatory, US Geological Survey, Hawaii 96718

Large earthquakes often occur as multiple ruptures reflecting strong variations of stress level along faults. Dense instrument networks with which the volcano Kilauea is monitored provided detailed data on changes of seismic velocity, strain accumulation and earthquake occurrence rate before the 1975 Hawaii 7.2-mag earthquake. During the ~4 yr of preparation time the mainshock source volume had separated into crustal volumes of high stress levels embedded in a larger low-stress volume, showing respectively high- and low-stress precursory anomalies.

OBSERVATIONS of seismic velocity anomalies by Russian seismologists[1,2] and the development of the dilatancy–diffusion model of earthquake precursors[3,4], have encouraged research into earthquake prediction. The simplicity of the microcrack–pore fluid mechanism[5], its foundation in repeatable observations from laboratory rock mechanics and the absence of contradictory evidence led many seismologists to anticipate in vain that definite prediction of earthquakes would soon be possible. For many medium-size earthquakes ($5<M<6$), mostly strike-slip events along the San Andreas fault, the expected precursors could not be detected[7–9]. These negative findings bring into question the universal applicability of dilatancy to the Earth's crust.

A new model of the preparatory process leading to earthquake ruptures is needed. However, field data for such a model cannot be gathered rapidly and in abundance because we do not know where and when large earthquakes will occur. Without an earthquake prediction capability we may have to wait several decades for large shocks to occur in a location where we can measure in detail the phenomena associated with their preparation process.

The Hawaii earthquake of 1975 occurred within a dense network of geophysical observations. Geodetic and seismological measurements had been carried on systematically for many years to monitor the volcano Kilauea. These data indicate that earthquake precursory changes do occur, but that they show contrasting behaviour in different parts of the earthquake source volume. We interpret this as meaning that the stress level throughout the source volume varies strongly—in small sub-volumes dilatancy-related precursory changes may occur, while strain-softening phenomena (due to stress unloading by creep) occur in other parts of the source volume.

The Hawaii (Kalapana) earthquake

The Hawaii earthquake of 29 November 1975 of 7.2 mag. occurred in the south-east coast of the island of Hawaii on Kilauea's south flank. The rupture occurred along a near-horizontal plane in the Earth's crust at a depth of 8 ± 1 km. Its length was 40–50 km and its width 10–20 km (Fig. 1). Along the south coast subsidence of up to 3.5 m occurred and tsunami heights reached 14.6 m (ref. 6). Damage from the earthquake and related catastrophic events was estimated at $4.1 million, and two people were killed.

The source of stress in Kilauea's south flank is the intrusion of magma into the volcano and its rifts[10,11]. The last earthquake of similar size to the 1975 rupture seems to have occurred in 1868. Since then repeated episodic intrusions of magma along the rifts (dotted in Fig. 1) have provided the 'pushing hand' compressing the south flank. The Hawaii earthquake was a rupture of the crust under stress, not a volcanic quake caused directly by an intrusion. While the stress loading process was different we believe that the rupture process was the same as that in other tectonic earthquakes.

Seismic waves emanating from this rupture were recorded by instruments on Hawaii which are designed to produce on-scale records for the strong ground motions of large, local earthquakes. The seismic energy was radiated in a sequence of high-amplitude pulses separated by comparatively long periods of low-amplitude radiation (Fig. 2). Such records are often observed for large earthquakes and are thought to represent multiple rupture sources[12–14]. As the rupture propagates along a heterogeneous faulting surface, or through a heterogeneous crust, it encounters high- and low-stress volumes. The rupture may speed up and slow down, or stop temporarily, with little radiation emitted from low-stress portions of the fault and large pulses emitted from zones of high stress. The latter may form around asperities on the fault surface, a term which we will use loosely as meaning a strong segment of the fault along which stresses will accumulate.

We conclude that the Hawaii earthquake source volume contained high- and low-stress sub-volumes, based on the strong

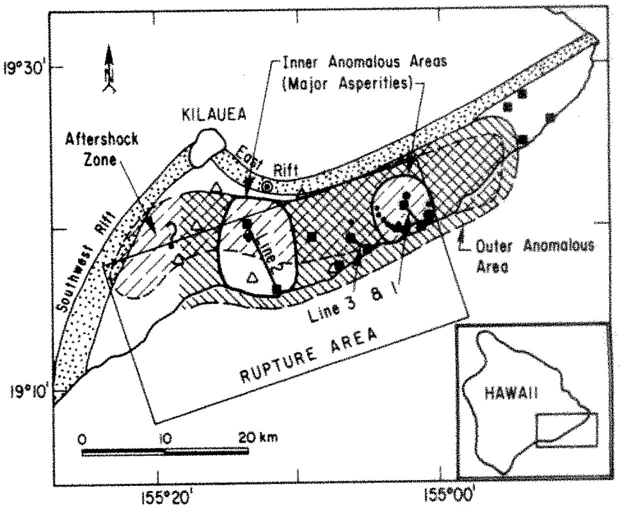


Fig. 1 Map of the south flank of Kilauea volcano showing major geological features, the aftershock zone of the Kalapana earthquake (dashed diagonal hatching), the outer precursor volume (solid hatching), the areas designated as 'major asperities', and the rupture area inferred from surface wave and tsunami data. △, Seismograph stations where constant seismic velocities were observed; ▲, the station showing the velocity anomaly; ■, geodetic line benchmarks; ⊙, the site of the 1972 Mauna Ulu eruption. ●, Immediate foreshocks; the 5.9-mag foreshock is the larger circle and the 7.2-mag main shock is the largest circle. Some seismic and geodetic stations are omitted for clarity.

0028–0836/81/040231—04$01.00

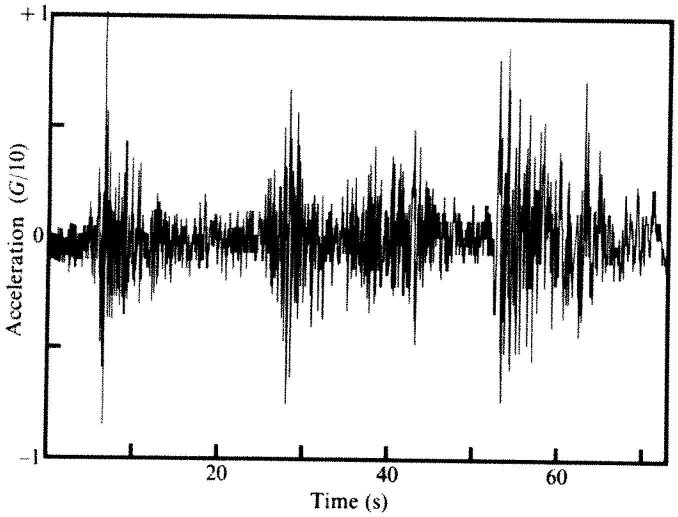


Fig. 2 Strong-motion seismogram of the Hawaii main shock of 29 November 1975, recorded at Hilo. Note that the main rupture seems to be made up of several events emanating energy pulses roughly comparable in amplitude and duration with that of the foreshock[22].

motion records (Fig. 2). It is therefore not surprising that in different regions of the source area we found contrasting precursory changes which we interpret as reflecting high- and low-stress regimes respectively.

Precursors to the Hawaii earthquake

The first documented precursor to the Hawaii earthquake was a delay of teleseismic P-wave travel times by 0.2 s lasting from 3.8 yr to several months before the main shock[15] (curve *b* of Fig. 6). According to Student's *t*-test this anomaly is significant at the 99.9% confidence level. Only the station nearest to the epicentre (▲ Fig. 1) exhibited delayed arrivals; six other stations (△, Fig. 1) within 3 km of the rupture area showed no significant change. Nearly horizontal rays from local earthquake sources also showed constant travel times at all stations, indicating that the anomalous volume did not extend to the surface[16].

These observations indicate that the Hawaii earthquake was preceded by a velocity decrease in a volume of ~10-km horizontal extent centred on the hypocentre, while the rest of the source volume remained nondilatant.

The rate of earthquake occurrence may contain information on the state of stress of a crustal volume. If the seismicity rate changes significantly as a function of time, the stress or the strength of the rock in the volume must have changed. Since 1968 all shocks with magnitude 2.0 or more are located routinely. During the 8 yr before the main shock ~1,600 such earthquakes occurred within the main rupture volume. This abundance enabled us to examine separately the rate in sub-volumes of roughly 10-km dimensions. Two types of seismicity behaviour were found. In some volumes the seismicity rate remained approximately constant. An example of this consistency is the seismicity rate of the foreshock volume (eastern major asperity in Fig. 1) as shown in curve 2 in Fig. 3. In other sub-volumes the seismicity rate dropped dramatically around the year 1971 ± 1 (see curve 1 in Fig. 3).

The combined seismicity of the volumes with approximately constant rate (inner anomalous areas of Fig. 1) contrasts sharply with that in the outer anomalous region (Fig. 4) where the rate dropped in early 1972 by ~50% from 120 events yr^{-1} to 55 events yr^{-1}.

The normal rate in the outer anomalous region is defined by data for the period 1968–71. A longer observation period can be opened if we raise the minimum magnitude in the data set. The seismicity rate in the outer anomalous region for magnitudes 2.4 or more is shown in Fig. 3 (curve 3) for the period 1962–75. Again we see that a highly significant decrease by about 50% occurred. The time of the change in this data set seems to be late 1970. We conclude that a 50% seismicity rate reduction occurred in the outer anomalous volume in 1971 ± 1 yr and lasted up to the time of the main shock.

Summarizing the geodetic data from 1914 to 1971, Swanson *et al.*[11] showed that the south flank was displaced several metres to the south-east with respect to observation points to the north and west of Kilauea's rifts. They also noted that the south flank accumulated compressive strain in the NW-SE direction in steps which coincided with volcanic eruptions and intrusions (Fig. 5). They concluded that shallow (depth <5 km), forceful magma intrusions under Kilauea and its rifts provided the 'pushing hand' which strained and dislocated the south flank away from Kilauea and the rest of the island. Based on these geodetic data, Swanson *et al.*[11] wrote that a large earthquake might be expected on the south flank.

All geodetic observation lines on the south flank showed compression up to 1970. In late 1971 and late 1974 respectively some lines reversed their behaviour, showing extension, that is they indicated strain was released[17]. The data for the past eight years are shown in more detail in Fig. 6*a*. Open circles indicate the average strain registered by lines 1 and 3 (defined in Fig. 1); solid squares show the strain measured on line 2. The strain data interpretation after 1972 is debatable because of the absence of measurements until late 1974. During this period a major phase of the east rift Mauna Ulu eruption commenced (large arrow in Fig. 6; ⦿ in Fig. 1). If the associated intrusive activity of this extrusive event compressed the south flank as observed for previous eruptions[11], it could account for the approximate 30-bar increase of line 2 observed in late 1974. Because no data are available for 1972–74, we show in Fig. 6 the uncertainty of the strain (stress) behaviour of this period with both a solid (preferred) and dotted line, depicting two alternative strain histories.

After the late-1974 survey, only line 2 in the western inner region was resurveyed before the 1975 main shock. It revealed a massive, pre-main shock strain relaxation, corresponding to a stress drop of ~30 bar. The earthquakes which occurred during this time could not account for this stress release because they were far too small. We conclude that the strain was released aseismically by some kind of creep process, perhaps along the future fault plane, and at a rate $>10^{-12}\ \mathrm{s}^{-1}$. Significantly, line 2, which accumulated stress up to the end of 1974, was located on a volume of constant seismicity rate, supporting the idea that high seismicity indicates high stress accumulation.

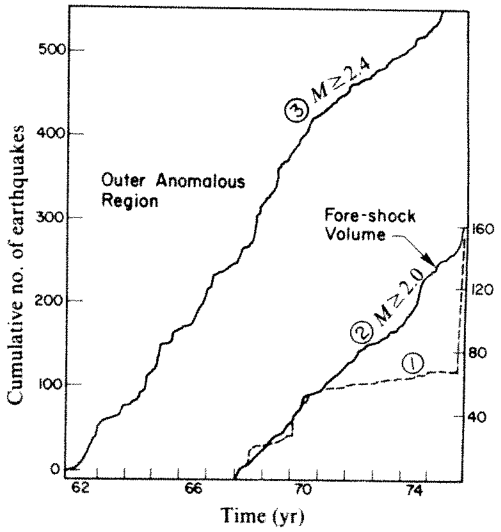


Fig. 3 Cumulative number of earthquakes as a function of time for three subregions of Kilauea's south flank. The slope of the curves equals the seismicity rate. The lower two curves (scale to the right) starting in 1968 are based on earthquakes with magnitudes >1.9. Curve 1 represents the south flank volume east of longitude 154.9° E, curve 2 the foreshock volume. For a better definition of the normal seismicity rate, the magnitude >2.3 seismicity from 1962–75 is shown as curve 3 (scale to the left) for the entire, outer anomalous region as defined in Fig. 1.

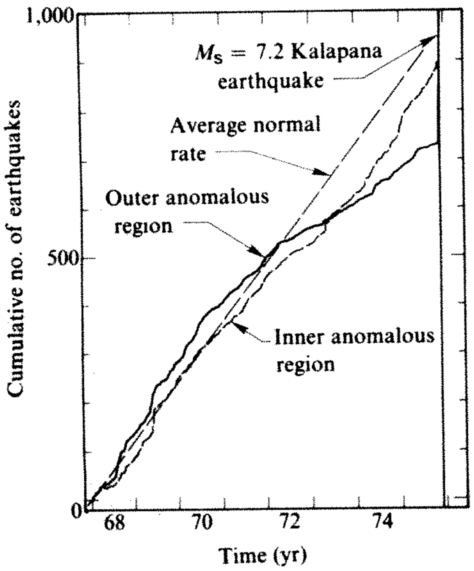


Fig. 4 Cumulative number of earthquakes ($M \geq 2.0$) in the south flank of Kilauea volcano, Hawaii, as a function of time. The solid curve, representing seismicity in the outer precursor volume (see Fig. 1), shows a sharp (≈50%) decrease in 1972. The dashed curve indicates roughly constant seismicity rate in the inner precursor volumes.

The strain release along line 1 (1971) and line 2 (1974) is interpreted to indicate that the respective southeastern end points of these lines were then located above the outer anomalous volume. Therefore the outer anomalous volume was apparently expanding at the expense of the inner anomalous volume.

The asperity model

The model for the preparatory process for the Kalapana earthquake must be complex, because different precursory patterns existed in subregions of the future rupture area. The characteristics of the outer anomalous zone are: decreased seismicity rate, strain relaxation and the absence of a travel time anomaly. One or both of the eastern (E) or western (W) inner anomalous zones, exhibited: constant seismicity rate until shortly before the main shock when the rate increased (E and W), a significant P-wave velocity decrease (E), the largest surface displacements during the main shock[6] (W) and an increase of compressive stress at least up to the fall of 1974 (W).

The decrease of seismicity in the outer volume seems to be directly related to the stress decrease observed there geodetically. Because the microearthquake activity in this region could not have accumulated the 10-bar stress drop throughout the outer anomalous volume, we hypothesize that a seismic fault creep along the nearly horizontal fault plane at 9-km depth caused the strain relaxation and the relative seismic quiescence. The stress decrease also explains the absence of travel time anomalies in the outer volume. The two inner anomalous volumes both have a constant and high seismicity rate. As this contrasts with the observations in the strain-softening outer anomalous volume, we suggest that the high seismicity rate indicates high stress in the inner anomalous zones. This interpretation is supported by different observations in the eastern and western volumes. In the eastern volume, near the point of initial rupture, the travel time anomaly indicates the presence of dilatancy, which requires high, deviatoric stresses. In the western volume, the stress increase of 30 bar between 1971 and 1974 shows that elevated stresses were present at that time. Thus, our observations indicate that two areas of high stress accumulation exist surrounded by a larger area where aseismic slip is occurring on the future rupture plane. By analogy with rock mechanics studies, we term these small volumes of high-stress concentrations 'major asperities'.

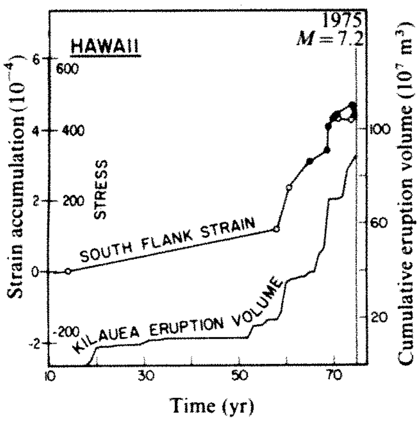


Fig. 5 Strain accumulation on Kilauea's south flank from 1914 to 1975 (upper curve). ○, Averages from the geodetic line 1 and 2 (Fig. 1); ●, strain along line 2. The cumulative eruption volume is shown for comparison. The strain direction is approximately southeast and the loading cycle is assumed to have started in 1868 after the last large earthquake on the south flank.

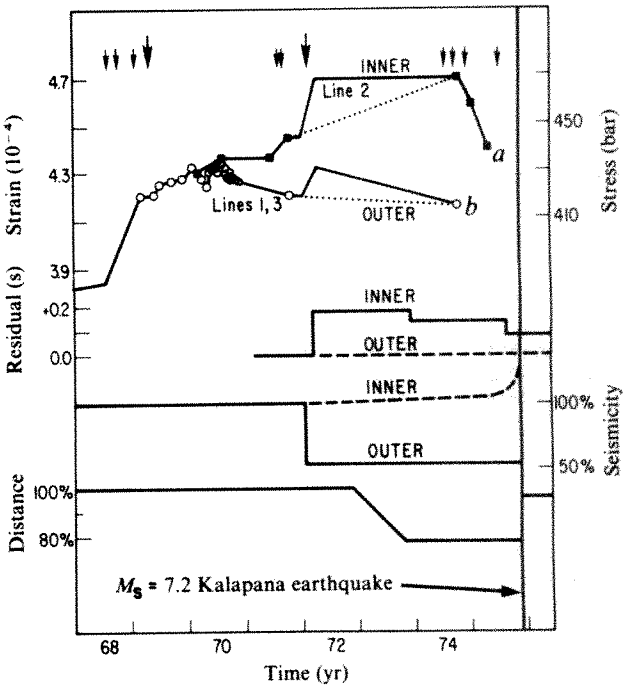


Fig. 6 Schematic summary of crustal parameters during the 8 yr before the Kalapana earthquake. Arrows at the top indicate times and relative size of volcanic intrusive episodes. The top curve shows geodetically measured horizontal strain increase in the western major asperity volume (■) and strain relaxation in the outer precursor volume (○). The dotted lines represent a possible alternate strain history between 1972 and 1975 when no measurements were made. The second curve shows the seismic-wave travel-time anomaly observed in the eastern asperity and the absence of such an anomaly in the outer precursor volume. The third curve shows the seismicity rates in the inner and outer precursor volumes (from Fig. 3). The final graph illustrates that the (normalized) average distance of earthquakes from the location of the large foreshock decreased beginning in 1972, which is an alternate way of quantitatively measuring the contrast between seismic activity in the inner and outer anomalous areas. Note that the anomaly onsets of the independent parameters occur at approximately the same time 1971 ± 1 yr. This strongly supports the idea that all of these events are related to the main-shock preparatory process, which may have been initiated by the 1972 large eruption/intrusion event.

The geodetic line 1 within the eastern asperity showed extension (relaxation) starting in 1970, and the line within the western asperity extended in early 1975 (Figs 5 and 6). This may mean that these two lines were contained entirely within their asperities until 1970 and 1975 respectively. After these times the creeping portions of the fault plane extended under the respective southern geodetic stations and the associated line began to lengthen. The central and northern parts of the asperities may have been locked until the time of the main shock. Unfortunately there is no geodetic data during the crucial period 1971–74, and the exact time of strain reversal is unknown.

This model of highly stressed asperities embedded in a strain-softening region of lower stress was derived from precursory geophysical data accumulated from the rupture zone of the Kalapana earthquake. It could, however, be considered a corollary of the known multiple-event nature of the main rupture (Fig. 2). Because the main rupture indicates that high-stress volumes were mixed with low-stress ones, the state of stress before the main shock must also have been one of contrasting high- and low-stressed volumes.

Based on laboratory observations, models similar to ours including strain-softening[18] and hard inclusions within strain-softening surroundings[19,20] have been proposed.

Conclusions

If our model is an accurate representation of the preparatory process to major earthquakes, it may prove difficult to detect high-stress-related earthquake precursors because they may be developed only in small portions of the precursor volume. If our observations can be scaled linearly to smaller and larger earthquakes in other tectonic environments, we expect that for a 5-mag earthquake the major asperities would have radii of ~1.5 km which may explain why no clear velocity anomalies have been detected for such events. For great earthquakes of several hundred kilometre source length, the asperities might have radii of ~40 km, which means that high-stress precursors might be detected with dense observation networks. The relationships between the number of asperities and earthquake size is not known. Our data do not exclude the possibility that large earthquakes contain numerous asperities with radii smaller than 5 km.

The precursor time for the Kalapana earthquake agrees approximately with the precursor time versus magnitude relations proposed elsewhere[3,4,21]. We have no evidence for diffusion of pore fluids, so the precursor time will more likely be a function of the ratio of locked-fault area to creeping-fault surface.

We conclude that three independent precursors (Fig. 6) have been observed for the Hawaii earthquake, which constrain the anomaly onset to approximately 1971 ± 1yr. They are of two types: high-stress precursors (decrease in P-wave velocity and foreshocks) and by contrast low-stress precursors, here observed as a 50% decrease in seismicity and strain relaxation. The high-stress phenomena are restricted to small volumes embedded in a larger volume characterized by lower stresses. The main-shock nucleation point was located in one of the major asperities and the area of largest surface deformations in the other. Our precursor model of hard asperities surrounded by creeping-fault area is analogous to the multiple rupture model for earthquakes. We suggest that future earthquakes may be predicted if patterns similar to those described here are observed.

We thank Robert Koyanagi and the staff of the Hawaiian Volcano Observatory for collecting and organizing data. This work was supported by NSF grant EAR 780-3633.

Received 9 June; accepted 26 September 1980.

1. Savarensky, E. F. *Tectonophysics* **6,** 17–27 (1968).
2. Semenov, A. N. *Bull. Acad. Sci. USSR, Phys. Solid Earth* **3,** 245–248 (1969).
3. Whitcomb, J. H., Garmany, J. D. & Anderson, D. L. *Science* **180,** 632–635 (1973).
4. Scholz, C. H., Sykes, L. R. & Aggarwal, Y. P. *Science* **181,** 803–810 (1973).
5. Nur, A. *Bull. seism. Soc. Am.* **62,** 1217–1222 (1972).
6. Tilling, R. I. *et al. U.S. Geol. Surv. Circ.* No. 740, **33,** (1976).
7. Kanamori, H. & Fuis, G. *Bull. seism. Soc. Am.* **66,** 2017–2037 (1976).
8. Bolt, B. A. *Bull. seism. Soc. Am.* **67,** 27–32 (1977).
9. Lindh, A. G., Lockner, D. A. & Lee, W. H. K. *Bull. seism. Soc. Am.* **68,** 721–734 (1978).
10. Ando, M. *J. geophys. Res.* **84,** 7616–7626 (1979).
11. Swanson, D. A., Duffield, W. A. & Fiske, R. S. *U.S. Geol. Surv. Profess. Pap.* 963 (1976).
12. Wyss, M. & Brune, J. N. *Bull. seism. Soc. Am.* **57,** 1017–1023 (1967).
13. Trifunac, M. D. & Brune, J. N. *Bull. seism. Soc. Am.* **60,** 137–160 (1970).
14. Kanamori, H. & Stewart, G. S. *J. geophys. Res.* **83,** 3427–3434 (1978).
15. Johnston, A. C. *Science* **199,** 882–885 (1978).
16. Johnston, A. C. thesis, Univ. Colorado (1979).
17. Lipman, P. W., Lockwood, J. P., Okamura, R. T., Swanson, D. A. & Yamashita, K. M. *U.S. Geol. Surv. Profess. Pap.* (in the press).
18. Mjachkin, V. I., Brace, W. F., Sobolev, G. A. & Dieterich, J. H. *Pure appl. Geophys.* **113,** 169–181 (1975).
19. Brady, B. T. *Pure appl. Geophys.* **112,** 701–725 (1974).
20. Stewart, W. D. *Tectonophysics* **52,** 613–626 (1979).
21. Rikitake, T. *Devl. Solid Earth Geophys.* **9,** 357 (1976).
22. Rojahn, C. & Morill, B. J. *Bull. seism. Soc. Am.* **67,** 493–515 (1977).
23. Wyss, M., Klein, F. W. & Johnston, A. C. *J. geophys. Res.* (in the press).

Upper mantle seismic anisotropy and lithospheric decoupling

J. H. Leven, I. Jackson & A. E. Ringwood

Research School of Earth Sciences, Australian National University, PO Box 4, Canberra 2600, Australia

The lithosphere underlying stable continental shield regions is thought to extend to a depth of ~200 km. Compressional wave velocity models derived from a recent approximately NS refraction profile in northern Australia include a pronounced high-velocity zone at a depth of around 200–250 km. Conventional petrological interpretations in terms of mineralogical or compositional changes cannot readily explain this feature. Calculated velocities for the garnet pyrolite model (using a derived geotherm), fit the velocities above and below this high-velocity zone, but cannot account for the high-velocity (~8.6 km s^{-1}) zone itself. This feature is explained in terms of velocity anisotropy, which might reflect mechanical decoupling of the continental lithosphere from the underlying mantle.

IT has been suggested that the lithosphere underlying continental shield regions extends to depths of the order of 200 km, and that the lithosphere (tectosphere)[1–3] beneath stable continental shields is depleted in low melting point components, is cooler than the normal mantle material at this depth, and hence would translate with the continental plates by virtue of its higher viscosity. Body and surface wave studies of the upper mantle structure beneath shield regions have not indicated any significant low-velocity zone above a depth of 200 km. A detailed study of the shear wave velocity structure beneath the Australian shield[4] has not found evidence for any region of low shear wave clocity above a depth of 150 km.

Lehmann[5,6] first postulated the existence of a seismic discontinuity in the upper mantle at a depth of around 200 km. This feature has subsequently been observed in many upper mantle refraction surveys, commonly at depths shallower than 200 km.

0028-0836/81/040234—06$01.00

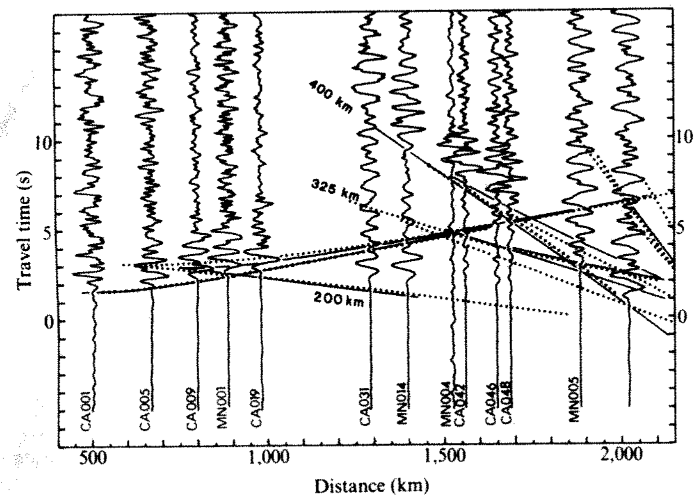


Fig. 1 Vertical component record section of event 256 from the Banda Sea Arc, with the travel time–distance curves of the model 8 (ref. 9) (dotted line) and CAPRI[11] (solid line) velocity models superimposed. The branch labelled 200 km corresponds to rays refracted below the 200 km discontinuity. This branch displays a rapid decrease in amplitude beyond an epicentral distance of 1,400 km, corresponding to the onset of a negative velocity gradient beneath the 200 km discontinuity (see Fig. 2).

Anderson[7] named this upper mantle feature the Lehman discontinuity. Hales[8] suggested that the relative motion of the Australian continental lithosphere over the underlying mantle material occurs "below a discontinuity in the P-wave velocity at a depth of about 200 km". Continents can move with comparable velocities to oceanic plates, so where is this zone of relative motion of the continental lithosphere over the underlying mantle and what is the relation, if any, of the '200 km' discontinuity to this zone?

Seismic observations

Hales *et al.*[9] analysed data from a long range upper mantle refraction profile conducted in Central Australia. Their upper mantle velocity model (model 8) has a velocity discontinuity at a depth of 200 km, with an increase in the compressional-wave velocity of 0.38 km s^{-1}. The depth of this feature has been constrained by the recorded arrivals of an event on the 14 October 1977 (event no. 256). The record (Fig. 1, see also Fig. 14 of ref. 9) of this event, which has been located in the depth range of 176 ± 5 km (on the basis of 164 readings), indicates that the 200 km discontinuity is significantly below the 176 km depth. The method of Gutenberg[9,10] constrains the velocity at the source depth of this event to be $\leq$8.28 km s^{-1}. Beneath the 200 km discontinuity, the velocity is well constrained to be 8.62 km s^{-1}, indicating that the velocity increase across the 200 km discontinuity is $>$0.3 km s^{-1}.

The 200 km prograde refraction branch, corresponding to rays which bottom below the 200 km discontinuity, shows an abrupt decrease in amplitude at an epicentral distance of around 1,700 km (depending on the source depth). Figure 12 of ref. 9 shows a fit of the travel time–distance curve derived from their model 5, to the record section for event no. 256. Model 5 did not incorporate any discontinuous increase in velocity around a depth of 200 km, so that this figure indicates the ease with which the 200 km discontinuity can remain undetected, especially when only shallow events are studied. Hales *et al.*[9] modelled the 200 km structure as a velocity knee with a layer having a negative velocity gradient below. Leven[11], using synthetic seismograms to model amplitudes, showed that the amplitude decay of the 200 km prograde branch can be explained by a low velocity zone which starts more sharply, around 30 km beneath the 200 km discontinuity. The exact nature of the velocity profile within the low velocity zone cannot be determined from present refraction data. However, synthetic seismogram modelling indicates that the negative velocity gradient commencing at a depth of 230 km must be sharper than that of model 8 (ref. 9), to explain this rapid decrease in amplitude near 1,700 km distance.

A sharp velocity increase at a depth of around 200 km, followed beneath by a velocity decrease, is a peculiar seismological feature. Similar high-velocity zones in the velocity models of other upper mantle studies can be found—but usually at shallower depths. For example, Lewis and Meyer[12] have a comparable structure at a depth of around 130 km, while in the model of Wiggins and Helmberger[13], this structure is at a depth of around 150 km. The abrupt decrease in amplitude of the 200 km branch has also been observed in other upper mantle refraction profiles, although the low velocity zone which is implied by this decrease has not been incorporated in the models resulting from these studies. Examples of this abrupt decay in the amplitude of the prograde 200 km branch can be clearly seen in Fig. 8 of ref. 14 and in Figs 11 and 12 of ref. 15. However, no convincing evidence of this type of velocity structure has been observed in the data of any of the Early Rise experiment profiles, which traverse North America with an excellent coverage of azimuth[16].

The existence of a negative velocity gradient at around 200 km depth is supported by the S to P conversion evident in records from teleseismic events. Sacks *et al.*[17] studied these converted phases recorded at the NORSAR array, and have interpreted precursors to the S onset as being S to P conversions from a negative velocity gradient at a depth of 250 ± 15 km.

Long range refraction surveys for oceanic paths have also indicated a similar velocity structure, but at a considerably shallower depth. Both Asada and Shimamura[18] and Hales *et al.*[19] have an analogous feature in the velocity models derived from oceanic refraction surveys. These velocity models both have a discontinuous increase to a compressional velocity of around 8.6 km s^{-1}, immediately above a low-velocity zone, the depth of this high-velocity zone in both cases being around 75 km. Leeds and others[20,21] have derived a model for the variation of the lithospheric thickness over the Pacific basin from surface wave studies. Their model indicates a thickening of the oceanic lithosphere from around 76 km in the younger regions, to 104 km in the older regions of the oceanic basin.

Conventional interpretations

Traditionally, seismic discontinuities in the Earth's mantle have been explained in terms of isochemical changes in mineralogy, changes in chemical composition, or a combination of both. The possible relevance of such phenomena to the 200 km discontinuity is now reviewed.

(1) Exsolution of $M_3^{2+}Al_2Si_3O_{12}$ (M = Mg, Fe, Ca) garnet from aluminous pyroxenes. The pressure-sensitivity of the Al_2O_3 content of pyroxenes coexisting with garnet provides the basis of pyroxene geobarometry[22,23]. Along the 1,200 °C isotherm, for example, the Al_2O_3 content of orthopyroxene coexisting with garnet in both synthetic and natural systems decreases, rather rapidly from 5–7% near 2.5 GPa to 1.5–2.5% near 4 GPa, and then more gradually to ~0.7% at 7 GPa and 0.4% in Ca-poor clinopyroxene at 14 GPa [ref. 23]. Green and Ringwood[24] (see also ref. 25) have shown that the orthopyroxene Al_2O_3 isopleths in the garnet pyrolite stability field are subparallel to the Clark and Ringwood[26] Precambrian shield geotherm—suggesting approximate constancy (at the 1.0–1.5% level) of orthopyroxene alumina content over the 2–5 GPa pressure range. The low alumina content of such orthopyroxene and the evidence for its gradual reduction with increasing pressure beyond 5 GPa exclude any possibility that exsolution of pyrope-rich $M_3^{2+}Al_2Si_3O_{12}$ (M = Mg, Fe, Ca) garnet from aluminous orthopyroxene is responsible for the discontinuous increase in compressional wave velocity near 200 km depth.

(2) Formation of complex $M_3^{2+}\,{}^{VI}(Al_2, M^{2+}Si)Si_3O_{12}$ garnet solid solutions. Ringwood's[27] study of the crystallization of a glass of 90% $MgSiO_3$–10% Al_2O_3 composition revealed a rapid

increase in both the garnet lattice parameter and the proportion of garnet near 10 GPa. More recently Akaogi and Akimoto[28] have confirmed the high pressure solubility of $MSiO_3$ pyroxene in the garnet structure, but suggest that such solution occurs gradually over the pressure range 4–14 GPa, with a pronounced increase in the proportion of garnet (and hence the density and P-wave velocity) in the 14–19 GPa pressure interval. The fact that the garnet lherzolite nodules of deepest origin (~200 km depth, 7 GPa pressure) brought to the Earth's surface by the Lesotho[23], Kao[22] and other kimberlite pipes contain 3.02–3.07 silicon atoms per 12 oxygen atoms is marginal evidence for the initiation of $M_3^{2+}(Al_2)Si_3O_{12} - M_3^{2+}(M^{2+}Si)\,Si_3O_{12}$ solid solution near this depth. However, neither laboratory experiments nor the garnet-bearing nodules indicate any abrupt increase in the $M_3^{2+}(M^{2+}Si)Si_3O_{12}$ content of garnet near 200 km depth in the Earth's mantle. This, coupled with the absence of other mineralogical changes in garnet lherzolite in the relevant pressure range, suggests that phase transformations are not responsible for the rapid increase in P-wave velocity near 200 km depth in the sub-continental mantle.

(3) Compositional change: harzburgite to garnet lherzolite. The long and complex geochemical history of subcontinental uppermost mantle, the geochemistry of ultramafic nodules from kimberlite pipes, and the absence of significant seismic low-velocity zones beneath Precambrian shields suggest that the sub-continental uppermost mantle is likely to be strongly (but variably) stripped of its low melting-point fraction. Under these circumstances, a transition would be expected, perhaps near 200 km depth, from barren harzburgite above, to fertile garnet lherzolite below[1,25]. The densities and P-wave velocities for these assemblages have been estimated to be 3.31 g cm^{-3} and 8.32 km s^{-1} for harzburgite and 3.38 g cm^{-3} and 8.38 km s^{-1} for pyrolite[25]. Jordan[29] has shown that removal (by 20% partial melting) of the basaltic fraction of pyrolite, decreases the density of the residium by ~0.05 g cm^{-3}, but leaves the P-wave velocity almost unchanged. The small change in P-wave velocity accompanying such differentiation is due to the near-cancellation between the velocity decrease associated with garnet removal and the velocity increase associated with the more magnesian composition of the residual olivine and orthopyroxene[29].

Clearly, although the subcontinental upper mantle must be chemically zoned as a result of differentiation, it is virtually impossible that a P-wave velocity discontinuity of the order of 0.3 km s^{-1} near 200 km depth, can be explained in these terms.

(4) Compositional change: garnet lherzolite (or harzburgite) to eclogite. Anderson[7,30] has suggested that the Lehmann discontinuity might represent a change in chemical composition from garnet lherzolite (or harzburgite) above the discontinuity to eclogite below. Such eclogite, whether produced by subduction of oceanic basaltic crust (as in Anderson's model), or by the necessarily relatively slow crystallization of basaltic magmas within the mantle, would be expected to possess an average mid-ocean ridge basaltic Mg number ($100\,Mg/(Mg+Fe^{2+})$) of about 60 (see refs 31, 32). In the former case, the basaltic major element chemistry might be perturbed during subduction by partial melting in the quartz eclogite stability field, resulting principally in significantly increased $(MgO+FeO)/CaO$ in the residium (see ref. 33). Typical mantle-derived eclogites are quartz-free, essentially bimineralic aggregates of pyroxene and garnet (in approximately equal proportions) with Mg numbers around 60.

Mean compressional-wave velocities for such mantle-derived eclogites[34–36] (samples 11, 14, and 15 of ref. 36 but not 8 which is very magnesian (Mg number 71) and contains 74.0% clinopyroxene, 16.6% garnet, and 7.9% zoisite) span the range 8.27–8.50 km s^{-1} at 1.0 GPa with measured velocity anisotropies of 0.8–3.4%. Shear-wave velocities are more variable (4.49–4.86 km s^{-1} at 1.0 GPa) with anisotropies of 1.5–4.9%. The Nové Dvory (Czechoslovakia) eclogite studied by Christensen to 3.0 GPa is representative of the highest P-wave velocity eclogites with $\rho = 3.56$ g cm^{-3}, and extrapolated compressional and shear wave velocities of 8.33 and 4.59 at atmospheric pressure. Our preferred parameters for garnet pyrolite are $\rho = 3.40$ g cm^{-3}, $V_p = 8.31$ km s^{-1}, and $V_s =$ 4.82 km s^{-1} (these differ slightly from those of Jordan[29] because of our use of the pyrope elastic moduli measured by Leitner *et al.*[37]).

Our analysis thus strongly reinforces other observations[7,25,34] concerning the indistinguishability of garnet lherzolite and eclogite P-wave velocities. We conclude that a change in chemical composition from garnet lherzolite to eclogite might produce a significant decrease in shear wave velocity but that no discernable change in compressional wave velocity would be expected. Under these circumstances, it seems unlikely that the Lehmann discontinuity reflects a chemical change of this type.

(5) Base of pronounced low-velocity zone. Many seismic models for upper mantle velocities beneath tectonically active continental regions (such as the western US) include a substantial and rapid increase in seismic wave velocities in the depth range 150–200 km (refs 38, 39). In most such models, the discontinuity represents the base of a pronounced low-velocity zone, so that a large part of the velocity contrast across the discontinuity is accounted for by a return to 'normal' upper mantle velocities (for example, at substantially subsolidus temperatures) below the discontinuity. The absence of significant low-velocity zones, shallower than 200 km depth, beneath Precambrian shield regions precludes such an interpretation for the Lehmann discontinuity.

Seismic-wave velocities for garnet pyrolite

The above arguments demonstrate the difficulty of accounting for the observed 0.3 km s^{-1} increase in V_p at ~200 km in terms of conventional explanations. To understand the nature of the seismic feature near 200 km depth, we now compare the CAPRI velocity-depth model[11] with calculated P-wave velocities for an assumed garnet pyrolite upper mantle.

For this calculation, we assume that the compressional-wave velocity (V_p of 8.31 ± 0.05 km s^{-1} at STP) varies linearly with pressure (to ~10 GPa) and temperature (to ~1,400 °C). Because volumetric strains will be no greater than 8% at 10 GPa, finite strain equations of state should not be necessary. The general tendency for both pressure and temperature derivatives of the elastic-wave velocity to decrease with increasing pressure and temperature suggests that linear extrapolation, using derivatives measured near STP, may result in slight overestimation of velocities at the upper end of the extrapolation. Any relaxation of ultrasonically determined shear moduli for upper mantle seismic-wave propagation would similarly result in an overestimation of the seismic velocities.

The pressure and temperature derivatives, used in the calculation:

$$\partial V_p/\partial P = 0.106 \text{ km s}^{-1}\text{ GPa}^{-1}$$

$$\partial V_p/\partial T = -0.55 \times 10^{-3} \text{ km s}^{-1}\text{ K}^{-1}$$

represent a weighted average (57% olivine, 17% orthopyroxene, 12% clinopyroxene, 14% garnet) of the mineral pressure and temperature derivatives. Data for Stillwater bronzitite ($\partial V_p/\partial P = 0.137$ km s^{-1} GPa^{-1} (refs 35, 40), are used in preference to the much higher value $\partial V_p/\partial P =$ 0.207 km s^{-1} GPa^{-1} derived from a study of single-crystal bronzite[41]. The latter value is grossly inconsistent with the expectation[42] that $(1/V_p)(\partial V_p/\partial P)$ is approximately equal to

$$\frac{39K + 16\,\mu}{6K(3K + 4\,\mu)}$$

This relation yields an expected $\partial V_p/\partial P$ of 0.138 km s^{-1} GPa^{-1}, in agreement with the measured value for Stillwater bronzitite. Furthermore, even if the unusually large single-crystal pressure derivatives ($dK/dP = 9.6$) were, in fact, appropriate at STP, it seems improbable that such a high derivative would be maintained over a substantial pressure range. Use of Frisillo and

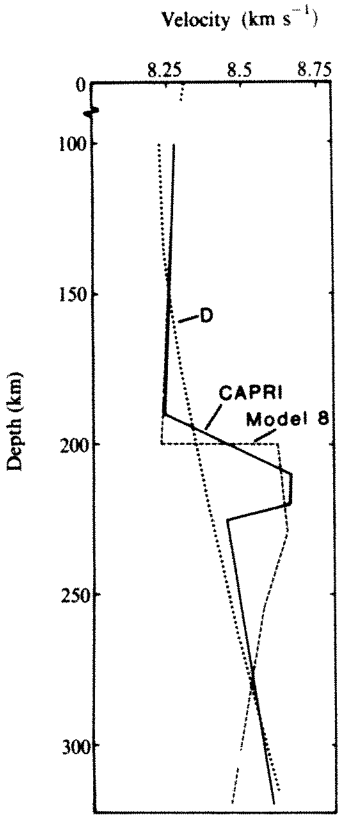


Fig. 2 A comparison of the model 8 and CAPRI compressional-wave velocity–depth models with calculated velocities (*D*) for a garnet pyrolite upper mantle based on the geotherm presented in the text.

Barsch data for orthopyroxene[9] would increase the garnet pyrolite $\partial V_p/\partial P$ to 0.128 km s^{-1} GPa^{-1} and hence raise the velocity at 200 km depth by ~0.14 km s^{-1}. Further experimental work on single-crystal and polycrystalline orthopyroxene is needed to resolve this uncertainty. Olivine and garnet pressure and temperature derivatives were taken from refs 43–45. A clinopyroxene ($\partial V_p/\partial P$) of 0.119 km s^{-1} GPa^{-1} was obtained by comparing Christensen's[35] data for Nové Dvory eclogite with the Bonczar *et al.*[45] data for garnet. Otherwise, it was assumed that the orthopyroxene derivatives also apply to clinopyroxene. In the initial calculations we used the geotherm of Harte[46]. The resulting compressional-wave velocity profile was broadly consistent with the model 8 of Hales *et al.*[9] and the CAPRI model in the 75–190 km and 225–320 km depth ranges but failed to explain the high velocities in the 190–225 km interval. Variation of the geotherm within reasonable limits[29] allowed a more satisfactory fit (Fig. 2) to the CAPRI velocities in the 75–190 km and 225–320 km depth ranges. The required geotherm

depth (km): 100; 150; 200; 250; 300
temperature (°C): 770; 1,050; 1,180; 1,320; 1,420

is between, and qualitatively similar to, Jordan's[29] S1 and S2 conductive geotherms, and would intersect adiabats for the oceanic upper mantle at depths between 300 and 400 km.

The comparison presented in Fig. 2 implies that the garnet pyrolite composition and mineralogy, combined with a reasonable geotherm, explains compressional-wave velocities for all but a narrow interval (190–225 km) of the 75–300 km depth range. Similar difficulties in explaining high velocities just below 200 km are evident in the studies of Graham[47], and of Ahrens[48] who noted that none of his calculated models "predict the 8.65 km s^{-1} P-velocity maximum . . . at 200 km depth in profile HWNE" (of Wiggins and Helmberger[13]). There have been similar comments[49] on the difficulty of explaining compressional-wave velocities $\geqslant$8.5 km s^{-1} in some seismic models of the oceanic lithosphere, in terms of the calculated elastic properties of petrologically appropriate isotropic aggregates.

The match between the calculated velocities and the CAPRI model, other than in the 190–225 km depth range, precludes neither a more barren harzburgitic chemistry above 200 km, nor a gradual increase in the $M_3^{2+}(M^{2+}Si)Si_3O_{12}$ component of the garnet with increasing depth below 200–250 km.

Seismic velocity anisotropy?

We are clearly concerned, not with a seismic discontinuity in the usual sense, but rather with a zone of the order of 50 km thick, characterized by higher velocities than both overlying and underlying mantle.

We have demonstrated the difficulty involved in explaining the 0.3 km s^{-1} velocity increase near 200 km, in terms of mineralogical or chemical changes. The negative velocity gradient below 230 km is also not readily explained in conventional terms. Leven[11] inferred that this velocity gradient must be more negative than −0.002 km^{-1} s^{-1} km^{-1} to produce synthetic seismograms consistent with the observed rapid decrease in first arrival amplitudes near $\Delta = 1{,}700$ km. Using P-wave velocity pressure and temperature derivatives appropriate for garnet pyrolite, we find that an unusually large (but not impossible) temperature gradient, of at least 9 °C km^{-1}, would be required if the negative velocity gradient below 230 km depth[9], were to be explained in terms of the combined effects of pressure and temperature.

On the other hand, this sub-continental 'high-velocity zone' bears a strong qualitative resemblance to shallower high-velocity lids capping low-velocity zones in many velocity models for the sub-oceanic upper mantle. These considerations, and the variability of its observation suggest that this unusual seismic feature might be explained in terms of seismic velocity anisotropy associated with deformation-induced preferred orientation of olivine (and possibly orthopyroxene) crystals in a shear zone, at 200–250 km depth, which is responsible for decoupling of the continental lithosphere from the underlying mantle. This hypothesis is supported by the following observations.

(1) Extensive field and laboratory studies of deformed peridotites from the Lanzo massif in Northern Italy (believed to represent subcontinental mantle), have established a firm relationship between the observed 7% velocity anisotropy, the fabric and the flow coordinates[50,51]. The velocity anisotropy is associated with preferred orientation of the [100] axes of olivine crystals at a low angle of inclination ($\theta = 10°$, Fig. 3) to the flow direction during high temperature (~1,200 °C) deformation at an inferred strain rate of 10^{-14} s^{-1}.

(2) Experiments conducted over a wide range of temperature and strain rate[51–54], have demonstrated the development of strong preferred orientation of olivine ([010]//σ_1 and [100]//σ_3) as a consequence of laboratory deformation of dunite. At high temperatures and moderate strain rates (moderate flow stress), the preferred orientation seems to be derived primarily from syntectonic recrystallization driven by strain accumulation, due principally to dislocation glide on the system (010)[100]. The coincidence of principal axes of stress and strain in most laboratory deformation experiments[52,54] results in an ambiguity concerning predicted fabric orientation in natural triaxial deformations which may resolve the apparent conflict between the orientations of the natural and laboratory fabrics mentioned above.

(3) Studies of the azimuthal variation of compressional-wave velocities in the oceanic upper mantle[55,56] have established the velocity anisotropy of the sub-oceanic mantle. The fast direction in each case is sub-parallel to the spreading direction, that is normal to the mid-ocean ridge. The azimuthal range in P-wave velocities (~0.3 km s^{-1}) and the association of fast velocities with the direction of plate movement over underlying mantle are both appropriate for explanation of the high P-wave velocities at 200–250 km depth along a NS profile beneath the Australian plate (Fig. 3).

(4) The seismic implications of the development of such an olivine preferred orientation are more fully explored in Fig. 3. For small inclinations of the fast [100] axis to the flow direction (N), propagation directions within 45° of the flow direction display compressional (or quasi-compressional) mode velocities significantly higher than 8.42 km s^{-1} which is the Voigt–Reuss–Hill (VRH) average. Note that the considerable shear wave birefringence—the difference in quasi-shear mode velocities is as much as 0.7 km s^{-1} for some propagation directions. Thorough characterization[57] of an essentially monominerallic sample of Twin Sisters dunite with a strong preferred orientation revealed a high degree of seismic anisotropy ($7.98 \leq V_p \leq 9.24$ km s^{-1}, $4.57 \leq V_s \leq 5.01$ km s^{-1}) and shear wave birefringence (maximum $\delta V_s = 0.38$ km s^{-1}). In upper mantle garnet peridotite, the magnitude of these effects will be reduced because of the reduced proportion of olivine. Nevertheless, the results of Fig. 3 provide a firm semiquantitative basis for testing the preferred orientation hypothesis.

The model provides a qualitative explanation of the high compressional-wave velocities for approximately NS propagation at depths of 200–250 km in the Australian upper mantle. The results presented in Fig. 3 (see especially $\theta = \phi = 22.5°$) suggest that significant shear wave birefringence might also be observed (on both horizontal and vertical components) at appropriate azimuths.

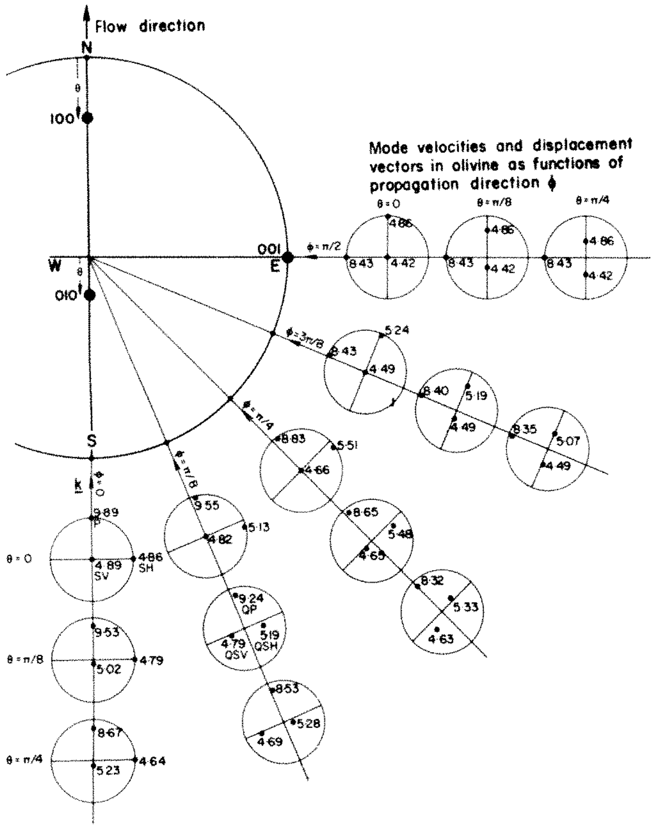


Fig. 3 Seismic anisotropy and shear-wave birefringence of olivine. The crystal orientation is the expected preferred orientation of olivine (see text) in the Australian upper mantle: flow plane horizontal, plate movement over underlying mantle due North. The olivine *c*-axis (001) lies EW while the inclination θ of the *a*-axis (001) to the flow direction is varied between 0 and $\pi/4$ (see text). Mode velocities and displacement vectors for horizontal propagation vectors **k** are calculated as functions of azimuth ϕ and inclination θ from the single crystal moduli[43]. The particle displacement vectors for the three independent propagation modes for each (ϕ, θ) pair are plotted (●) on a stereographic projection and labelled with the appropriate velocities (in km s^{-1}). Note that off-axis propagation directions involve modes for which particle displacement is neither parallel nor orthogonal to the propagation vector. Such coupled modes are labelled QP, OSH, QSV[67]. The results may be applied in the SW quadrant by virtue of the mirror symmetry of olivine about the plane normal to 001.

(5) Both the chemistries and textures of deep-seated ultramafic modules contained in kimberlite pipes are consistent with the view that decoupling of the continental lithosphere and the underlying mantle is concentrated at depths near 200 km (ref. 58). The observation that ultramafic nodules of deepest origin (~200 km), brought to the surface by South African kimberlites varying in age by more than 1,000 Myr (refs 22, 59) display more primitive chemistry than their shallower counterparts, suggests that the former are derived from a zone of the mantle which is continuously replenished from a large mantle reservoir. Furthermore, while the highly sheared mosaic textures of the deepest-seated xenoliths are probably not representative of steady-state mantle deformation processes[60], the contrasting fabrics of these and the relatively undeformed coarse-grained xenoliths[49] of shallower origin, are suggestive, at least, of very different modes of xenolith incorporation and magma transport.

(6) The combination of more primitive chemistry below 200 km depth and the general characteristics of upper mantle geotherms suggests that the zone between 200 and 300 km depth may be characterized by higher homologous temperatures (T/T_m) for olivine than for the shallower and the deeper mantle. Under these circumstances, this region is a plausible zone for strain concentration, although the relative roles of dislocation creep—syntectonic recrystallization (which may be subject to strain hardening) and grain boundary sliding[61], remain to be determined. Anderson[30] similarly concludes that the viscosity of subcontinental mantle should attain a minimum value near 200–250 km depth.

Conclusions

We have demonstrated the difficulty involved in explaining high compressional-wave velocities for approximately NS propagation at 200–250 km depth in the Australian upper mantle in terms of various suggested variations of mantle mineralogy and chemistry. Having rejected conventional interpretations, we suggest that the observed high P-wave velocities may derive from preferred orientation of olivine crystals in a deformation zone responsible for the mechanical decoupling of the lithosphere from the underlying mantle.

Refraction profiles of the Australian upper mantle[9,62–64], while in general agreement on the need for a substantial increase in compressional-wave velocity near 200 km depth for profiles within 45° of NS, provide limited information for other azimuths. The only indication of a possible 200 km discontinuity on an approximately EW profile is provided by a single first arrival[65] at $\Delta \sim 2{,}100$ km (Meekatharra, Western Australia) from an explosion at Mt Fitton (South Australia). This datum, while possibly compatible with NS profiles, does not provide a strong independent constraint on the velocity below 200 km depth for EW propagation azimuths. Clearly, the anisotropy hypothesis needs further testing. Carefully controlled experiments over a wide range of azimuths (such as those conducted on the suboceanic mantle[56]) will be required to test for anisotropy and possible shear wave birefringence (see ref. 66) of the continental upper mantle.

However, observation of subcontinental upper mantle seismic anisotropy may be subject to considerable variability. Spatial variation in the abundances of key volatile species (such as H_2O), and of mantle geotherms and strain rates may result in variability of the dominant upper mantle deformation mechanism, with some plausible flow mechanisms (such as grain boundary sliding) having little tendency to produce a strong preferred orientation. Even where an appropriate mechanism (such as dislocation creep–syntectonic recrystallization) dominates, the development of an observable seismic anisotropy may depend on a long, consistent and current history of relative motion of the lithospheric plate with respect to underlying mantle.

Such anisotropy beneath both continental and oceanic lithosphere, would facilitate seismic mapping of the vertical extent of the lithosphere, and provide considerable insight into the rheological behaviour of the upper mantle.

We thank P. N. Chopra, D. A. Gust, A. L. Hales, and K. J. Muirhead for useful discussion.

Received 26 June; accepted 11 November 1980.

1. Ringwood, A. E. *J. geophys. Res.* **67,** 857–867 (1962).
2. Jordan, T. H. *Rev. Geophys. Space Phys.* **13,** 1–12 (1975).
3. Jordan, T. H. *Nature* **274,** 544–548 (1978).
4. Hales, A. L., Muirhead, K. J. & Rynn, J. M. W. *Geophys. J. R. astr. Soc.* (submitted).
5. Lehmann, I. *Ann. Geophys.* **15,** 93–118 (1959).
6. Lehmann, I. *Bull. seism. Soc. Am.* **52,** 519–526 (1962).
7. Anderson, D. L. *J. geophys. Res.* **84,** 7555–7560 (1979).
8. Hales, A. L. *Director's Report, Research School of Earth Sciences Annual Report* (Australian National University, Canberra, 1977).
9. Hales, A. L., Muirhead, K. J. & Rynn, J. M. W. *Tectonophysics* **63,** 309–348 (1980).
10. Gutenberg, B. *Bull. seism. Soc. Am.* **3,** 223–232 (1953).
11. Leven, J. H. thesis, Australian National University (1980).
12. Lewis, B. T. R. & Meyer, R. P. *Bull. seism. Soc. Am.* **58,** 565–596 (1968).
13. Wiggins, R. A. & Helmberger, D. V. *J. geophys. Res.* **78,** 1870–1880 (1973).
14. Masse, R. P., Landisman, M. & Jenkins, J. B. *Geophys. J. R. astr. Soc.* **30,** 19–36 (1972).
15. Johnson, L. R. *J. geophys. Res.* **72,** 6309–6325 (1967).
16. Warren, D. H. *et al. Project Early Rise, Travel Times and Amplitudes* (U.S. Geological Survey Open File Report, Menlo Park, 1968).
17. Sacks, I. S., Snoke, J. A. & Husebye, E. S. *Tectonophysics* **56,** 101–110 (1979).
18. Asada, T. & Shimamura, H. in *The Geophysics of the Pacific Ocean Basin and its Margin* (eds Sutton, G. H. *et al.*) 19 (American Geophysical Union, Washington DC, 1976).
19. Hales, A. L., Helsley, C. E. & Nation, J. B. *J. geophys. Res.* **75,** 7362–7381 (1970).
20. Leeds, A. R., Knopoff, L. & Kausel, E. G. *Science* **186,** 141–143 (1974).
21. Leeds, A. R. *Phys. Earth planet. Inter.* **11,** 61–64 (1975).
22. MacGregor, I. D. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A.) (American Geophysical Union, Washington DC, 1979).
23. Akaogi, M. & Akimoto, S. *Phys. Earth planet. Inter.* **19,** 31–51 (1979).
24. Green, D. H. & Ringwood, A. E. *Earth planet. Sci. Lett.* **3,** 151–160 (1967).
25. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
26. Clark, S. P. & Ringwood, A. E. *Rev. Geophys.* **2,** 35–88 (1964).
27. Ringwood, A. E. *Earth planet. Sci. Lett.* **2,** 255–263 (1967).
28. Akaogi, M. & Akimoto, S. *Phys. Earth planet. Inter.* **15,** 90–106 (1977).
29. Jordan, T. H. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A.) (American Geophysical Union, Washington DC, 1979).
30. Anderson, D. L. *Geophys. Res. Lett.* **6,** 433–436 (1979).
31. Cann, J. R. *Phil. trans. R. Soc.* **A268,** 495–505 (1971).
32. Melson, W. G., Vallier, T. L., Wright, T. L., Byerly, G. & Nelen, J. in *The Geophysics of the Pacific Basin and its Margin* (eds Sutton, G. H. *et al.*) (AGU Monogr. No. 19, 1976).
33. Stern, C. R. and Wyllie, P. J. *Am. Miner.* **63,** 641–663 (1978).
34. Kanamori, H. & Mizutani, H. *Bull. Earthquake Res. Instn.* **43,** 173–194 (1965).
35. Christensen, N. I. *J. geophys. Res.* **79,** 407–412 (1974).
36. Manghnani, M. H., Ramananantoandro, R. & Clark, S. P. *J. geophys. Res.* **79,** 5427–5446 (1974).
37. Leitner, B. J., Weidner, D. J. & Liebermann, R. C. *Phys. Earth planet. Inter.* **22,** 111–121 (1980).
38. Archambeau, C. B., Flinn, E. A. & Lambert, D. G. *J. geophys. Res.* **74,** 5825–5864 (1969).
39. Helmberger, D. V. & Wiggins, R. A. *J. geophys. Res.* **76,** 3229–3245 (1971).
40. Wang, C. *J. geophys. Res.* **79,** 771–772 (1974).
41. Frisillo, A. L. & Barsch, G. R. *J. geophys. Res.* **77,** 6360–6384 (1972).
42. Birch, F. in *The Earth's Crust and Upper Mantle* (ed. Hart, P. J.) 13 (American Geophysical Union, Washington DC, 1969).
43. Kumazawa, M. & Anderson, O. L. *J. geophys. Res.* **74,** 5961–5972 (1969).
44. Graham, E. K. & Barsch, G. R. *J. geophys. Res.* **74,** 5949–5960 (1969).
45. Bonczar, L. J., Graham, E. K. & Wang, H. *J. geophys. Res.* **82,** 2529–2534 (1977).
46. Harte, B. *Phil. Trans. R. Soc.* **A288,** 487–500 (1978).
47. Graham, E. K. *Geophys. J. R. astr. Soc.* **20,** 285–302 (1970).
48. Ahrens, T. J. *Phys. Earth planet. Inter.* **7,** 167–186 (1973).
49. Green, D. H. & Liebermann, R. C. *Tectonophysics* **32,** 61–92 (1976).
50. Peselnick, L., Nicolas, A. & Stevenson, P. R. *J. geophys. Res.* **79,** 1175–1182 (1974).
51. Nicolas, A. & Poirier, J. P. *Crystalline Plasticity and Solid-State Flow in Metamorphic Rocks* (Wiley-Interscience, London, 1976).
52. Ave'Lallemant, H. G. & Carter, N. L. *Bull. geol. Soc. Am.* **81,** 2203–2230 (1970).
53. Nicolas, A., Boudier, F. & Boullier, A. M. *Am. J. Sci.* **273,** 853–876 (1973).
55. Hess, H. *Nature* **203,** 629–631 (1964).
56. Raitt, R. W., Shor, G. G., Francis, T. J. F. & Morris, G. B. *J. geophys. Res.* **74,** 3095–3106 (1969).
57. Babuška, V. *J. geophys. Res.* **77,** 6955–6965 (1972).
58. Boyd, F. R. *Geochim. cosmochim. Acta* **37,** 2533–2546 (1973).
59. Danchin, R. V. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A.) (American Geophysical Union, Washington DC, 1979).
60. Goetze, C. *Geology* **3,** 172–173 (1975).
61. Chopra, P. N. & Paterson, M. S. *Tectonophysics* (in the press).
62. Muirhead, K. J., Cleary, J. R. & Finlayson, D. M. *Geophys. J. R. astr. Soc.* **48,** 509–519 (1977).
63. Denham, D., Simpson, D. W., Gregson, P. J. & Sutton, D. J., *Geophys. J. R. astr. Soc.* **28,** 225–235 (1972).
64. Simpson, D. W. thesis, Australian National Univ. (1973).
65. Finlayson, D. M., Cull, J. P. & Drummond, B. J. *J. geol. Soc. Aust.* **21,** 447–458 (1974).
66. Hirn, A. *Geophys. J. R. astr. Soc.* **49,** 49–58 (1977).
67. Crampin, S. *Geophys. J. R. astr. Soc.* **49,** 9–27 (1977).

Regulation of transcription in expressed and unexpressed mating type cassettes of yeast

Amar J. S. Klar, Jeffrey N. Strathern, James R. Broach* & James B. Hicks

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724
* Department of Microbiology, SUNY at Stony Brook, Stony Brook, New York 11790

The genes that control the a, α *and* a/α *cell types in* Saccharomyces *are carried on transposable elements known as* a *and* α *cassettes which reside at three different chromosomal loci. Examination of the transcripts by R-looping and filter hybridization indicates that each cassette is capable of producing two divergent transcripts. Cassettes at the* MAT *locus are transcribed constitutively. Transcription of cassettes at* HML *and* HMR *is prevented by* trans-*acting negative regulators.*

THE genes controlling mating type in *Saccharomyces* yeasts reside on movable elements, referred to as cassettes, which become activated on transposition from silent storage sites to the mating type locus (*MAT*)[1-5]. The particular mating type gene activated in a cell determines the mating type of that cell. Cells expressing only an *α* mating type gene are phenotypically *α* cells: that is, they can mate with *a* cells but not with other *α* cells. Conversely, cells expressing only an *a* mating type gene are phenotypically *a* cells: they can mate with *α* but not with *a* cells. Cells that express both *a* and *α* mating type genes display a third phenotype (designated *a*/*α*): they can mate with neither *a* nor *α* cells and, if diploid, can undergo meiosis and sporulation. All cells contain three distinct genetic loci on yeast chromosome III at which mating type genes may reside. At two of these loci, *HML* and *HMR*, the genes are normally unexpressed. Only the gene present at the *MAT* locus (either *a* or *α*) is expressed in wild-type cells.

We describe here observations on the transcription of mating type cassettes located at both the *MAT* locus and the *HML* and *HMR* storage sites in a variety of strains. The results of this study have allowed us to refine current hypotheses which deal with two specific questions of mating type regulation. First, what is the mechanism by which *MATa* and *MATα* products determine cell type? Second, what is the mechanism which allows expression of *a* and *α* DNA at *MAT* but keeps them unexpressed at *HML* and *HMR*?

Genetic analyses suggest that *MATa* and *MATα* do not code for all the cell-type specific functions, but determine cell type by regulating the expression of cell-type specific genes elsewhere in the genome[6]. Several of these structural genes involved in the sexual cycle have been identified as mutations that result in *a*- or *α*-specific defects in mating or as mutations defective in meiosis and sporulation of *a*/*α* cells. Two complementation groups in *MATα* and one complementation group in *MATa* have been identified and their roles in the regulation of *a*-specific, *α*-specific and *a*/*α*-specific genes defined[7-9]. *MATα*1 is a positive regulator of *α*-specific genes (*matα*1 mutants are sterile). *MATα*2 is a negative regulator of *a*-specific genes (*matα*2 mutants express both *a*-specific and *α*-specific genes and hence are sterile). *MATα*2 is also a positive regulator of *a*/*α*-specific

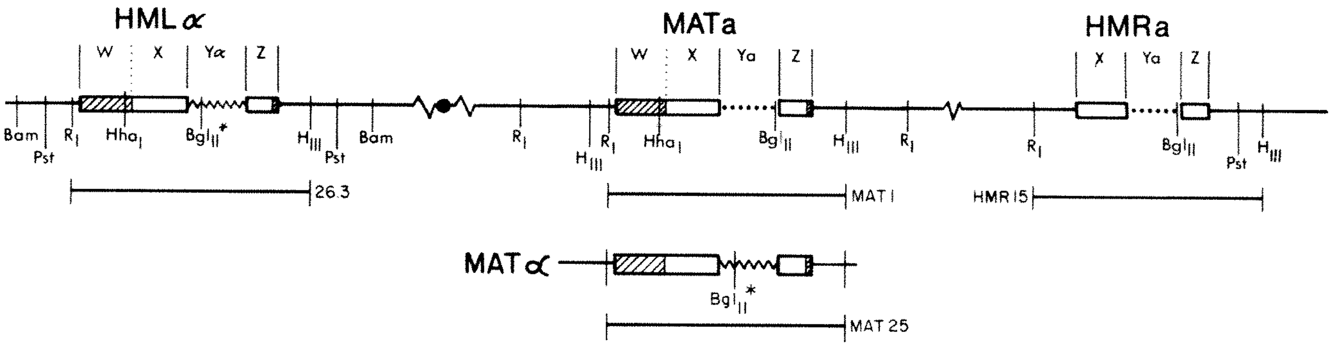


Fig. 1 Organization of the mating type cassettes on yeast chromosome III. Open boxes represent sequences common to all three loci, *MAT, HML* and *HMR*. Hatched regions are present at *HML* and *MAT* only. Sequences specific to *a* or *α* cassettes are denoted by dotted and zig-zag lines, respectively. The solid circle denotes the centromere. Detailed restriction maps are presented elsewhere[4]. Only restriction fragments used as probes in subsequent figures are labelled. The *BglII* site in the *α* cassettes labelled with an asterisk is presented only in strains derived from *Saccharomyces carlsbergensis*.

genes (*mataα2/MATa* diploids are sporulation defective). *MATa* has no role in the mating ability of *a* cells (*mata*$^-$ mutants mate efficiently with *α* cells), but *MATa* is required, along with *MATα*2, for the *a/α* phenotype (*mata*$^-$*/MATα* diploids behave like *α* cells)[10]. The observations reported here contribute further to our understanding of how *MATa* and *MATα* regulate cell type.

There is substantial genetic and structural information available regarding the mechanism by which the mating type cassettes at the storage sites are prevented from being expressed. The two storage sites for the *a* and *α* cassettes, *HML* and *HMR*, are located at opposite ends of the same chromosome in which the *MAT* locus resides[11]. Each storage site contains one copy of a mating type cassette, which can be either *a* or *α*[3-5]. Recent structural analysis of cloned DNA molecules from *MAT* and the storage loci, *HML* and *HMR*, has shown that all three loci share two regions of sequence homology flanking a 600- or 750-base pair sequence unique to *a* or *α*, respectively[3,4]. The physical structure of *HML, HMR* and *MAT* and their orientation on the chromosome are shown in Fig. 1. Finally, genetic analysis indicates that *HML* and *HMR* are under apparent negative control by the combined action of several distinct genes variously known as *MAR, SIR* and *CMT*[12-14]. These genes are genetically unlinked to each other or to *HML, HMR* or *MAT*: The *MAR*1 locus has been mapped to chromosome IV (ref. 12). *MAR*2 has been mapped to chromosome XII and is allelic to *CMT* (unpublished results). A mutation in either of these genes is sufficient to cause phenotypic expression of *HML* and *HMR*. These genes, therefore, operationally act as co-repressors of *HML* and *HMR*. In strains with the most common arrangement of storage genes, that is, *HMLα* and *HMRa*, the *mar*1-1 mutation yields a non-mating, *a/α*-like phenotype, presumably because both *α* and *a* functions are being expressed[12].

Thus, the mating type cassettes encode products which regulate expression of unlinked genes involved in the mating and sporulation pathways and are themselves regulated when located in the storage sites. To examine the nature of these controls further, we have examined the transcription patterns of the *MATa* and *MATα* genes in *a*, *α* and *a/α* cells. We have mapped these transcripts and assayed their expression in mutant straints defective in normal control functions. We have also examined transcription from the *a* and *α* elements residing at *HML* and *HMR* in normal strains and in strains carrying *mar*$^-$ mutations which result in the expression of these loci.

Identification of transcripts

The *MATa* and *MATα* alleles each produce two mature RNA transcripts. These transcripts were detected by electrophoretic separation of poly(A)$^+$-selected RNA on agarose gels followed by transfer to diazotized filter paper[15] and hybridization with ^{32}P-labelled *a* and *α* DNA cloned in bacterial plasmids. Figure 2 displays a series of RNA gel blots from normal haploid strains in which only the *MAT* locus is expressed and *mar*1$^-$ non-mating strains in which all three loci containing mating type cassettes are expressed (discussed below). Tracks *a* and *c* of Fig. 2 contain RNA from normal *MATa* and *MATα* haploid strains probed with an *Eco*RI–*Hin*dIII fragment of *HMLα* DNA, and tracks *e* and *g* contain the same RNAs probed with a *Eco*RI–*Hin*dIII fragment of *MATa* DNA. The sizes of the labelled transcripts were determined relative to pBR322 restriction fragment markers run on the same gel. A comparison of these four tracks yields the following conclusions: (1) there are two transcripts from *MATa*, one of 650 bases (*a*x) and one of approximately 450 bases (*a*y); (2) only the 650 base *MATa* transcript is homologous with *α* DNA from *HMLα* (Fig. 2, track *a*); (3) *MATα* produces a broad band (tracks *c*, *g*) which consists of two transcripts of similar size, ~800 bases (*α*1) and 750 bases (*α*2). An additional band high in gel tracks *e–h* represents a transcript not associated with mating type but which overlaps the cloned *MAT* fragment used as a hybridization probe. The *MATα* transcripts, which are not separated on the gel shown in Fig. 2, can be distinguished both by R-loop analysis in the electron microscope and by hybridization mapping, as described below. More detailed mapping of these transcripts is also presented below.

In addition to the normal haploid strains we have also analysed the patterns of transcripts in *mar*1$^-$ haploids. As described above, genetic characterization of strains harbouring the *mar*1$^-$ allele indicates that such strains express the storage loci, *HML* and *HMR*, in addition to the *MAT* locus[12]. Hybridizations to RNA from *mar*1$^-$ strains of genotype *HMLα MATα MATa* (Fig. 2, tracks *d, h*) and *HMLα MATa HMRa* (tracks *b, f*) are shown. In the tracks probed with *MATa* DNA both characteristic *a* transcripts appear as well as at least one *α* transcript. This pattern of bands is identical with that obtained using RNA from normal *a/α* diploid (see Fig. 4, track *j*). Therefore, we assume here that the state of mating type expression in non-mating *mar*$^-$ cells is equivalent to that in a *a/α* diploid. The relative intensities of the bands with tracks of RNA from *mar*1$^-$ strains correspond roughly to the gene dosage of *a* and *α* cassettes in each strain, suggesting that all three loci are being expressed at comparable levels. Comparison of *mar*1$^-$ and *MAR*1$^+$ strains indicates that the *HML* and *HMR* loci are apparently under transcriptional control through the *MAR* functions. Similar results have been obtained for mutants carrying the *mar*2$^-$ allele which defines a *MAR* function mapping on chromosome XII (data not shown). The tracks probed with *HMLα* yield a similar result with the exception that the small *a* transcript is missing, indicating that this 450-base transcript comes from the sequences unique to the *a* cassette.

Mapping the transcripts

The positions of both *α* transcripts (*α*1 and *α*2) and the longer of the two *a* transcripts (*a*x) were determined by electron microscopy of R-loops formed by hybridization of poly(A)$^+$ RNA to a *Eco*RI–*Hin*dIII fragment containing *MATa* or *MATα* DNA cloned in the bacterial plasmid pBR322. A histogram of

R-looped molecules is shown in Fig. 3. One α transcript was mapped to the X region of the *MAT* locus and the other extends from approximately the middle of the Yα region rightwards through Z. Both α transcripts would therefore be expected to show homology with the *MATa* gene as well as *MATα* because the X and Z regions are common to both (see Fig. 1). One of the *MATa* transcripts (ax) was also found to map in the X region of the *MAT* locus. It seems to share the same left-hand end point as the homologous α transcript ($\alpha 1$) but is somewhat shorter than $\alpha 1$, as expected from the gel blotting experiments. Only two molecules containing the a transcript (ay) were found in our R-loop preparations. However, blotting results in Fig. 2 and additional studies described below clearly indicate that ay is not homologous to the *MATα* gene and must therefore lie in the Ya region of the *MATa* gene.

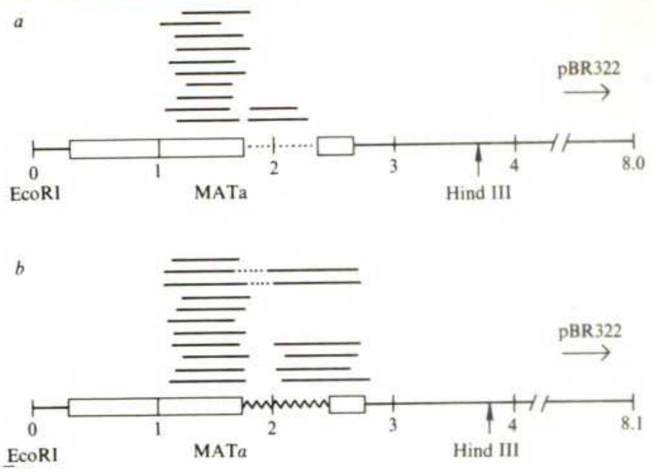


Fig. 3 Histogram of R-looped plasmid molecules containing the *MATa* gene (pBR-*MAT*1) hybridized with RNA from a *MATa* strain (a) and the *MATα* gene (pBR-*MAT*25) hybridized with RNA from a *MATα* strain (b). Both cloned *MAT* genes were inserted in pBR322 as *Eco*RI–*Hin*dIII fragments and the hybrid plasmids were digested with *Eco*RI. Dotted lines indicate individual molecules containing both R-loops, measured molecules represent approximately 50% of those observed to contain R-loops. Bacteriophage ΦX–double stranded DNA was included as a size standard. R-looping was performed according to the method of Kaback *et al.*[21] and measurements were made on a Numonics electronic planimeter.

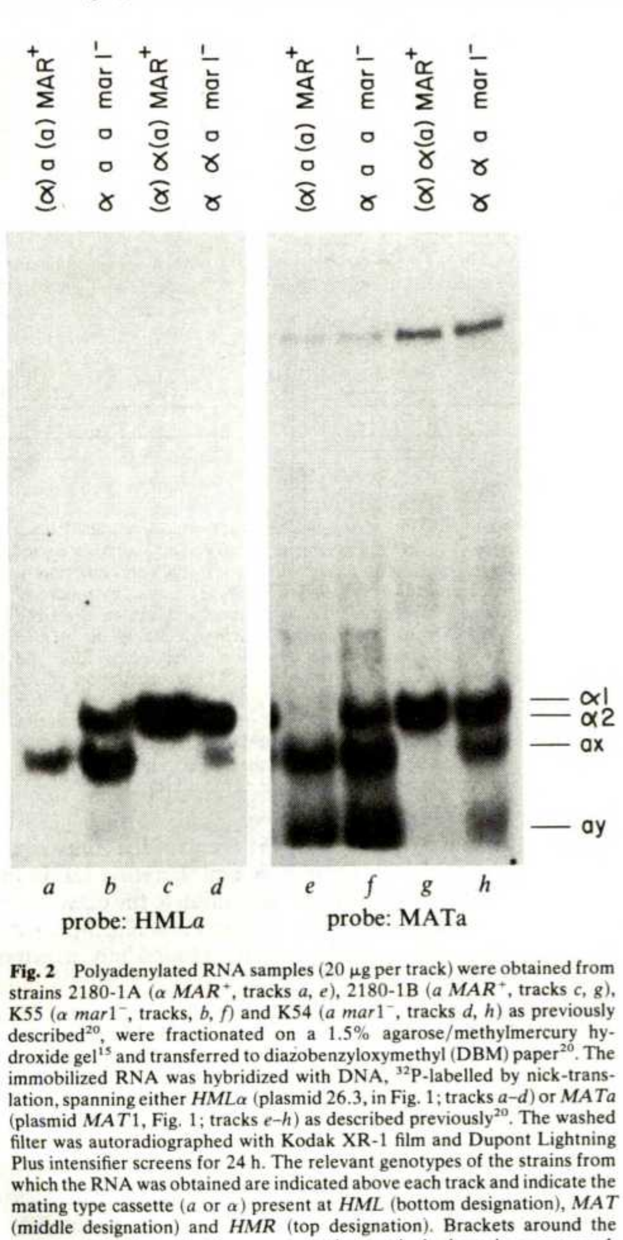


Fig. 2 Polyadenylated RNA samples (20 μg per track) were obtained from strains 2180-1A (α *MAR*$^+$, tracks a, e), 2180-1B (a *MAR*$^+$, tracks c, g), K55 (α *mar*1$^-$, tracks, b, f) and K54 (a *mar*1$^-$, tracks d, h) as previously described[20], were fractionated on a 1.5% agarose/methylmercury hydroxide gel[15] and transferred to diazobenzyloxymethyl (DBM) paper[20]. The immobilized RNA was hybridized with DNA, ^{32}P-labelled by nick-translation, spanning either *HMLα* (plasmid 26.3, in Fig. 1; tracks a–d) or *MATa* (plasmid *MAT*1, Fig. 1; tracks e–h) as described previously[20]. The washed filter was autoradiographed with Kodak XR-1 film and Dupont Lightning Plus intensifier screens for 24 h. The relevant genotypes of the strains from which the RNA was obtained are indicated above each track and indicate the mating type cassette (a or α) present at *HML* (bottom designation), *MAT* (middle designation) and *HMR* (top designation). Brackets around the designation indicate that the cassette at that particular locus is unexpressed, as established by genetic criteria. The absence of brackets indicates that the cassette is expressed at that particular locus. Bands corresponding to $\alpha 1$, $\alpha 2$, ax and ay transcripts, identified as described in the text, are indicated to the right of the autoradiogram. Molecular weights were estimated using restriction fragments of bacterial plasmid pBR322 as standards (not shown).

Additional mapping data were obtained by probing RNA blots with purified restriction fragments from cloned mating type genes as shown in Fig. 4. Tracks a–f in Fig. 4 represent a single pair of blotted tracks hybridized sequentially to three different probes. Tracks a, c and e contain RNA from *mar*1$^-$ non-mating a/α-like strain, tracks b, d and f contain RNA from a normal α haploid. Tracks a and b are probed with a fragment which covers Ya, Z and overlaps the X region of *MATa*. The probe reveals a single band in the a/α tracks, at ~750 bases, but shows a doublet at 800/750 bases in the α track. We interpret this result to mean that two transcripts are made from the *MATα* locus in strains exhibiting the α phenotype but only one in a/α strains. This interpretation was confirmed by additional hybridization studies described below. Knowing the positions of the α transcripts, we could confirm that only one of two α transcripts is present in a/α cells. This was done by testing the original filter in Fig. 4a,b with probes containing only X-region homology (103.4) and Z-region homology (103.5). Using the high molecular weight band associated with the right-hand end of the *MAT* probe as an internal standard it is clear (Fig. 4, tracks c–f) that the Z-region transcript ($\alpha 1$) is not present in the a/α strain but the X-region transcript ($\alpha 2$) is present in both α and a/α cells. The presence of two transcripts from *MATα* is consistent with the observation of two complementation groups *MATα*1 and *MATα*2 as described above. Further, the absence of the $\alpha 1$ transcript in a/α cells correlates with the fact that a functional *MATα*1 allele is not required for the a/α phenotype. On the other hand, *MATα*2 is a required function in a/α diploids. We therefore identify the message which is present only in α haploid strains ($\alpha 1$) as the transcript for the *MATα*1 gene and the message present in both α and a/α strains ($\alpha 2$) as the transcript of *MATα*2. This interpretation is consistent with genetic recombination results suggesting that *MATα*2 is proximal to *MATα*1 (refs 7, 9) and has been confirmed by *in vitro* mutagenesis by Nasmyth *et al.* (see accompanying article[16]).

Figure 4 also yields information concerning the location of the a transcripts. Both ax and ay transcripts are clearly present in a *HMLα MATα HMRa mar*1$^-$ strain when probed with whole *MATa* DNA (Fig. 2, track h; see also the a/α strain in Fig. 5). However, when probed with fragment 103.8, only the smaller of the a transcripts, ay, is revealed in the a/α track, indicating that the 650-base transcript (ax) lies mainly to the left of the *Hha*I restriction site in the region X. Furthermore, as the small a transcript does not hybridize significantly with *HMLα* DNA (Fig. 2), it must lie mainly within the Ya region. This interpretation is confirmed by additional hybridization of a/α and a strains with the X-region probe 103.7 (Fig. 4g,h) and an a strain with the Y-region probe 103.8 (Fig. 4i). In additional experiments (data not shown) the W region probe 103.6 was used against a and α RNA. No hybridization was detected with this probe to any of the RNA species. Because either mating type cassette can be expressed at *HMR*[17], even though the W region is not represented there[4,5], it was not required that those sequences be transcribed.

Direction of transcription

The direction of transcription of the mating type genes was determined by hybridization of RNA gel blots to separated strands of λ bacteriophage carrying cloned *HML*α or *HMRa* sequences followed by subsequent hybridization to wild-type λ DNA labelled with ^{32}P by nick translation. The orientation of the cloned genes in the λ vector was determined by heteroduplex analysis in the electron microscope. The result of these experiments is shown in Fig. 5. The tracks on the left represent parallel electrophoresis of poly(A)$^+$ RNA from an α haploid strain *S245c* hybridized with the R and L strand, respectively, of the hybrid λ clone *MAT*25 (*MAT*α). The $\alpha 1$ band hybridizes to the R strand of the λ clone, that is, it is transcribed from Yα towards Z, whereas $\alpha 2$ hybridizes to the L strand and is thus transcribed leftward as the locus is drawn in Figs 1 and 5. The pair of tracks on the right contain poly(A)$^+$ RNA from a a/α diploid strain probed with separated strands from λ clone *HMR*15c (*HMRa*). As described previously, this RNA preparation would be expected to contain the transcripts designated ax, ay and $\alpha 2$, all of which should be detected by the *HMR*15c probe. These tracks show that ax as well as $\alpha 2$ hybridizes to the L strand and is transcribed leftward in the X region of the cassette. The ay transcript hybridizes to the opposite strand and is thus transcribed to the right in region Ya. Thus, both a and α cassettes are transcribed divergently from within the middle of the cassette. At least for $\alpha 1$ and ay the 5′ end of the transcript is inside the unique Yα or Ya sequence and therefore inside the transposed portion of the cassette. In addition to the direction of transcription, a comparison of tracks a and c in Fig. 5 provides additional evidence that the $\alpha 1$ transcript is found only in cells of the α mating type and not in a/α diploids.

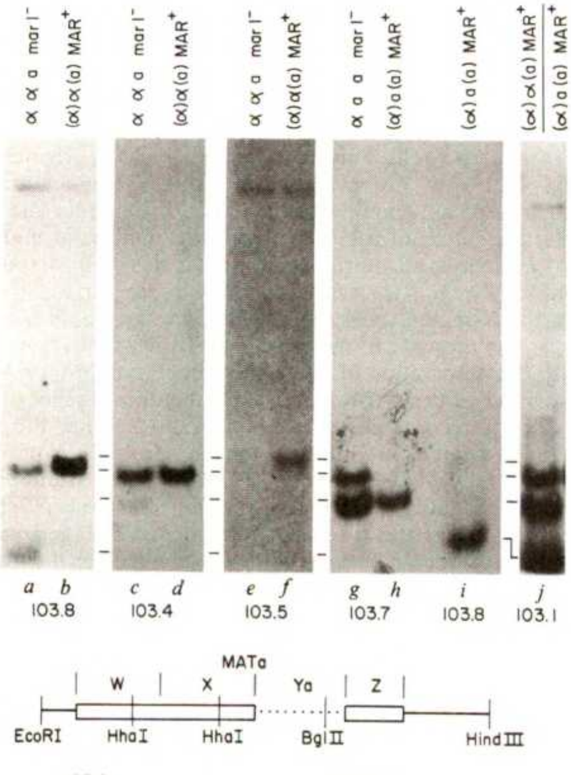


Fig. 4 Mapping the a and α transcripts by hybridization of RNA gel blots with specific probes. tracks a–f represent the same filter probed sequentially with the probes indicated. Between hybridizations the previous probe was removed by treatment with 99% formamide and 10 mM HEPES buffer, pH 7.5, at 65 °C for 1 h. Relevant genotypes of the strains used are indicated above each track as described in Fig. 2 legend. Parentheses denote silent *HML*α and *HMRa* cassettes.

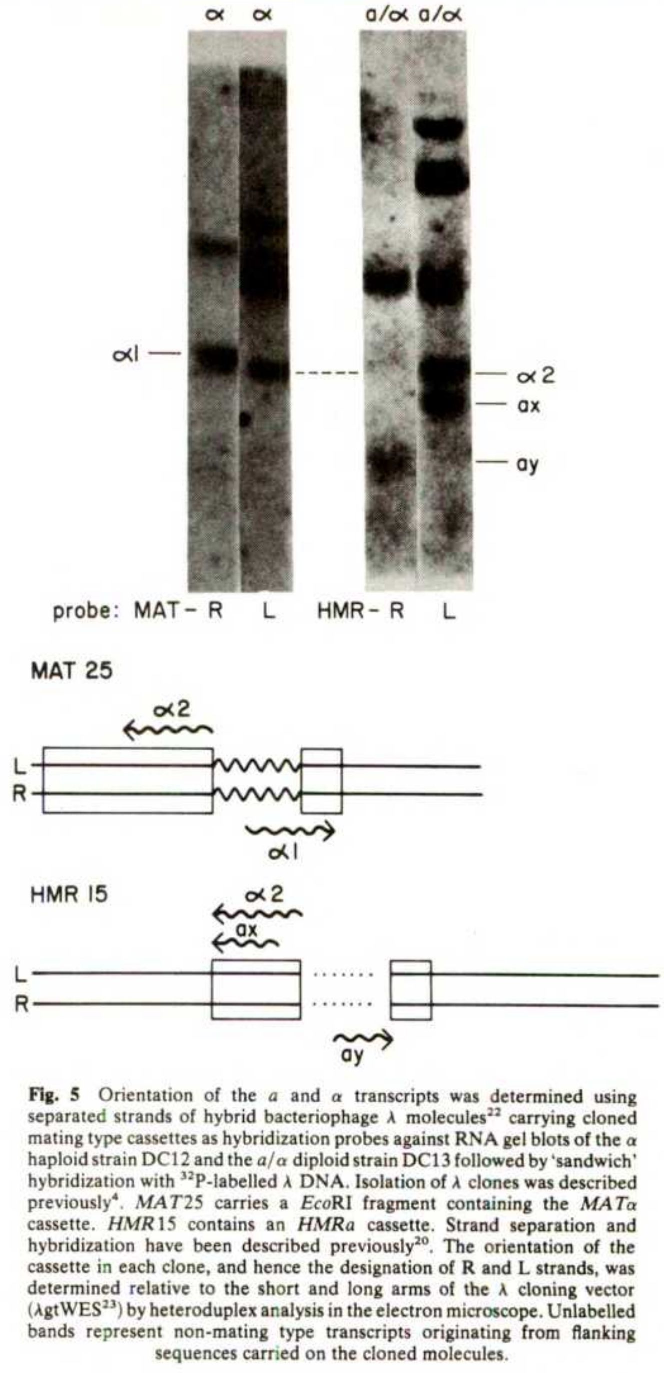


Fig. 5 Orientation of the a and α transcripts was determined using separated strands of hybrid bacteriophage λ molecules[22] carrying cloned mating type cassettes as hybridization probes against RNA gel blots of the α haploid strain DC12 and the a/α diploid strain DC13 followed by 'sandwich' hybridization with ^{32}P-labelled λ DNA. Isolation of λ clones was described previously[4]. *MAT*25 carries a *Eco*RI fragment containing the *MAT*α cassette. *HMR*15 contains an *HMRa* cassette. Strand separation and hybridization have been described previously[20]. The orientation of the cassette in each clone, and hence the designation of R and L strands, was determined relative to the short and long arms of the λ cloning vector (λgtWES[23]) by heteroduplex analysis in the electron microscope. Unlabelled bands represent non-mating type transcripts originating from flanking sequences carried on the cloned molecules.

Control of $\alpha 1$ transcription in a/α cells

The transcript mapping results described above showed that one of the two α transcripts ($\alpha 1$) is not present in a/α (non-mating) cells. This is consistent with genetic results indicating that *MAT*$\alpha 1$ is not required for expression of the a/α cell type but is required in α cells to turn on α-specific mating functions. We have therefore investigated whether *MAT*$\alpha 1$ is turned off by the products of the *MATa* allele (ax and/or ay) or by the combined action of *MATa* and *MAT*$\alpha 2$. Poly(A)$^+$ RNA was isolated from *mat*$\alpha 2$/*MATa* diploids and tested for expression of the $\alpha 1$ transcript by hybridization with the whole *MATa* probe (homologous to $\alpha 1$, $\alpha 2$, ax and ay) and a purified probe containing only the *MAT*α yz region specific to $\alpha 1$. The autoradiographs in Fig. 6 show sequential hybridizations of these two probes to the same filter containing two $mat\alpha 2^-/MATa$ diploids, a normal $MATa/MAT\alpha$ diploid and a $MAT\alpha$ haploid strain. The $\alpha 1$ and $\alpha 2$ bands are not resolved, but the positive signal at the $\alpha 1/\alpha 2$ position in the mutant diploids when probed with $MAT\alpha YZ$ indicates that the $\alpha 1$ transcript is present. As expected, no signal is seen at that position in the $MATa/MAT\alpha$ diploid. Thus, it is clear that both $MAT\alpha 2$ and the product(s) of the $MATa$ gene are required for turn-off of $MAT\alpha 1$. The possible mechanisms of $MAT\alpha 2$ and $MATa$ action are described further below.

Control of cell type

The results presented here, combined with genetic results described elsewhere, can be used to generate a circuit diagram for the control of cell type in *Saccharomyces*. Figure 7 shows the inferred regulatory pathways in α strains, *a* strains and a/α strains. In the α strain shown in Fig. 7*a* both $MAT\alpha 1$ and $MAT\alpha 2$ have active regulatory roles. Because $MAT\alpha 1$ is necessary for efficient mating but not for sporulation it is proposed to be a positive regulator of α-specific functions. $MAT\alpha 2$, on the other hand, is proposed to act as a negative regulator of *a*-specific functions. A $mat\alpha 2^-$ mutant is presumed to be defective in mating as both α and *a* mating functions are turned on.

In the mating type *a* strain shown in Fig. 7*b*, no control functions are assigned to the products of the $MATa$ allele. This is because no recessive mutations at the $MATa$ locus have been shown to cause defects in *a* mating ability. We presume, therefore, that the *a*-specific genes are constitutive in the absence of $MAT\alpha 2$, and that even though $MATa$ is transcribed in both *a* and a/α strains, its products only exert an effect in a/α strains.

Regulation in a $MATa/MAT\alpha$ diploid strain is shown in Fig. 7*c*. In addition to the turn on of a/α-specific functions by $MAT\alpha 2$ and $MATa$, some control mechanism shutting off *a*- and α-specific functions is required. Several genetic arguments suggest that $MAT\alpha 2$ is responsible for shutting off *a*-specific functions in a/α cells just as it does in α cells[9]. We show here that α-specific functions are turned off in a/α diploids by the combined action of $MAT\alpha 2$ and $MATa$ repressing the transcription of the positive regulator of α genes, $MAT\alpha 1$. This analysis does not determine whether it is *a*x or *a*y (or both) that are required for the a/α cell type. *In vitro* mutagenesis of cloned $MATa$-containing plasmids by Nasmyth *et al.* (see accompanying article[16]) indicates that only *a*y is necessary for this interaction. No function has been established for *a*x. We therefore propose that an interaction occurs between the $MAT\alpha 2$ and the $MATa$ product(s) which changes their regulatory specificities. This interaction creates a new regulatory molecule which serves to turn off $MAT\alpha 1$ and also to turn on the a/α-specific genes whose functions are required for meiosis and sporulation. Of course, it is also possible that $MAT\alpha 2$ and the products of $MATa$ do not interact with each other but must act at independent sites on the DNA so as to regulate $\alpha 1$ and the a/α specific genes.

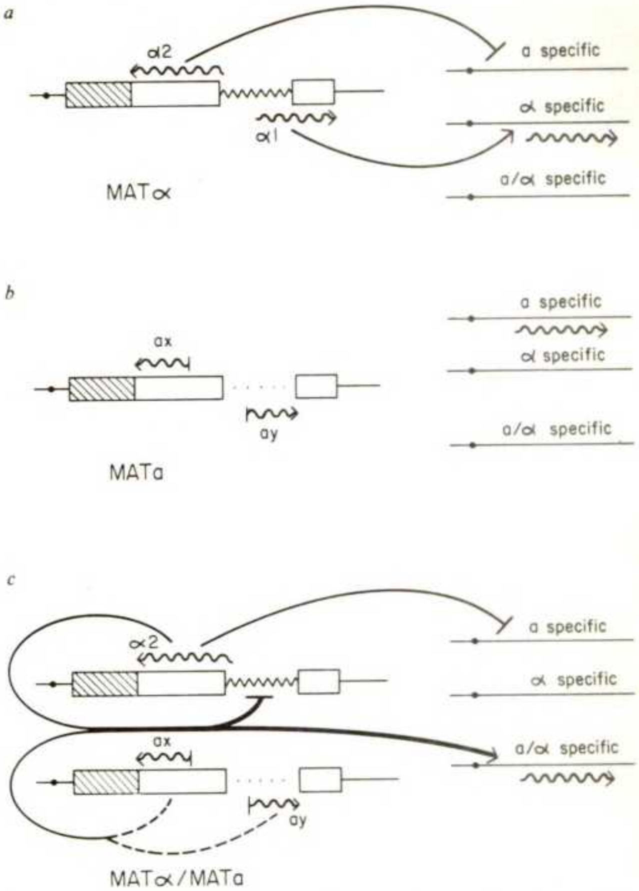


Fig. 7 Control of cell type in *Saccharomyces*. (*a*) *a* cells; (*b*) α cells; (*c*) a/α cells. Straight arrows indicate positive control. T-bars indicate negative control. Wavy arrows are the transcripts described in the text. Other symbols as in Fig. 1. The requirement of both *a*x and *a*y in control of cell type has not been determined. Genetic analyses of control of cell type are described elsewhere[6–10].

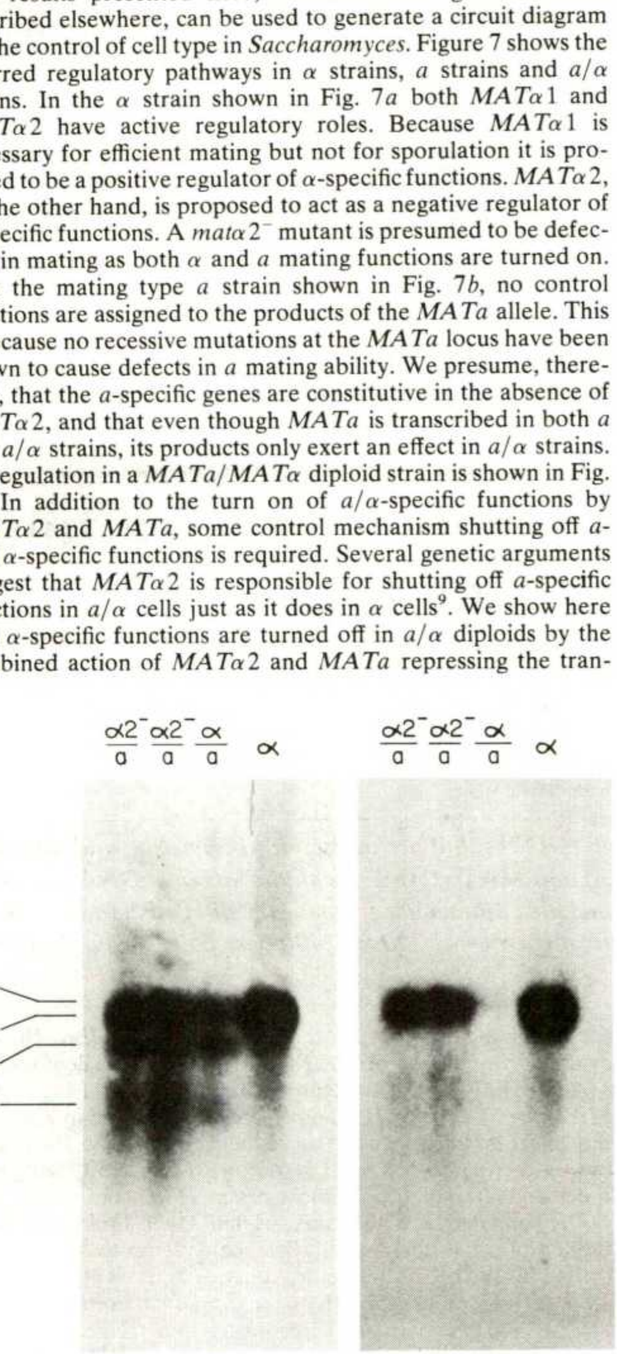


Fig. 6 Sequential hybridization of RNA from a series of strains blotted on diazotized paper[20] to probes containing a complete $MATa$ cassette ($MAT1$, Fig. 1) and the right half of a $MATa$ cassette ($MAT\alpha YZ$). Two independently isolated $mat\alpha 2^-/a$ diploids (DC100, DC101), a normal a/α diploid (DC13) and an α haploid (DC12) were used. The $MAT\alpha$ YZ probe was made by cutting a cloned α cassette from strain CB11 at the *Bgl*II site in the Y region[4] yielding a probe specific for the $mat\alpha 1$ transcript.

Control of the silent genes

In addition to the control of cell type itself the transcription mapping presented here provides insight into the mechanism by which the storage cassettes, *HML* and *HMR*, are kept silent. There is much genetic evidence indicating that *HML* and *HMR* are under negative control and the results presented here show that control is apparently at the level of transcription. At least four distinct genetic loci, known as *MAR* or *SIR*, have been identified which are required to prevent expression of *HML* and *HMR*. In both $mar1^-$ and $mar2^-$ mutants mature transcripts are produced at *HML* and *HMR* and these transcripts seem to be identical to their mature counterparts from *MAT*. These results are particularly surprising considering the 5′ end of at least one α transcript ($\alpha 1$) and one *a* transcript (*a*y) map within the region of the cassette which must be transposed during mating type switching.

How, then, can that region of the cassette be silent at one chromosomal locus and active at another? On one hand, it is possible that the 5′ end of the observed transcripts do not reflect the actual transcription initiation site. The transcripts could start near the ends of the cassettes and proceed in opposite directions, crossing each other to reach the regions maintained in the mature transcript. In this model the transcription initiation sites could be in or near regions not shared by *MAT* and the storage loci and hence subject to different control. On the other hand, transcriptional initiation may occur near the 5′ ends of the mature transcripts. In this case, differential regulation of the cassettes could be effected by differential accessibility of the promoter site depending on whether the cassette is located at *MAT* or at *HML* or *HMR*. The accessibility of the promoter would be established by the interaction of sites outside the regions of homology with the *MAR* and *SIR* products and could reflect differential chromatin structure[18], phasing of nucleosomes[19] or differential methylation. The structure of certain intrachromosomal rearrangements[4] suggests that the sites of action of the negative regulators lie to the left of the W–X regions at *HML* and *HMR*. Deletion of these sequences results in the expression of the normally silent *HML* and *HMR* cassettes.

Regardless of the particular mechanistic model, it is clear that mating type expression in yeast is an example of a genetic positional effect; that is, the expression of a DNA sequence depends on its location in the genome. The availability of mutants which allow expression of the silent cassettes and the existence of cloned DNA from *MAT*, *HML* and *HMR* should permit a detailed analysis of the molecular basis of this positional effect.

We thank David Kaback for assistance with the R-looping analysis, Ira Herskowitz for comments on the manuscript, Jean McIndoo, Carolyn McGill, Victoria Guarascio and Sajida Ismail for technical assistance, Louisa Dalessandro for preparation of the manuscript and Mike Ockler for the figures. This research was supported by NIH research grant GM25678-03 to A.K., GM25634-03 to J.H. and J.S and GM24226 to J.B.

Received 8 September; accepted 14 November 1980.

1. Hicks, J. B., Strathern, J. & Herskowitz, I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A., Shapiro, J. & Adhya, S.) 475–462 (Cold Spring Harbor, New York, 1977).
2. Harashima, S., Nogi, Y. & Oshima, Y. *Genetics* **77,** 639–650 (1974).
3. Hicks, J., Strathern, J. N. & Klar, A. J. S. *Nature* **282,** 478–482 (1979).
4. Strathern, J., Spatola, E., McGill, C. & Hicks, J. B. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2839–2843 (1980).
5. Nasmyth, K. A. & Tatchell, K. *Cell* **19,** 753–764 (1980).
6. McKay, V. & Manney, T. R. *Genetics* **76,** 273–288 (1974).
7. Strathern, J. N. thesis, Univ. Oregon (1977).
8. Hicks, J. B. thesis, Univ. Oregon (1975).
9. Strathern, J., Hicks, J. & Herskowitz, I. *J. molec. Biol.* (in the press).
10. Kassir, Y. & Simchen, G. *Genetics* **82,** 187–206 (1976).
11. Harashima, S. & Oshima, Y. *Genetics* **84,** 437–451 (1976).
12. Klar, A. J. S., Fogel, S. & Macleod, K. *Genetics* **93,** 37–50 (1979).
13. Rine, J., Strathern, J. N., Hicks, J. B. & Herskowitz, I. *Genetics* **93,** 877–901 (1979).
14. Haber, J. E. & George, J. P. *Genetics* **93,** 13–32 (1979).
15. Alwine, J. C., Kemp, D. J. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
16. Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. & Smith, M. *Nature* **289,** 244–250 (1981).
17. Klar, A. J. S., McIndoo, J., Hicks, J. B. & Strathern, J. N. *Genetics* (in the press).
18. Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R. & Elgin, S. *Cell* **16,** 797–806 (1979).
19. Gottesfeld, J. & Bloomer, L. S. *Cell* **21,** 751–760 (1980).
20. Broach, J. R., Atkins, J. F., McGill, G. & Chow, L. *Cell* **16,** 827–839 (1979).
21. Kaback, D. B., Angerer, L. M. & Davidson, N. *Nucleic Acids Res.* **6,** 2499–2517 (1979).
22. Szybalski, W., Kubinski, H., Hradecna, Z. & Summers, W. C. *Meth. Enzym.* **21,** 383–413 (1971).
23. Leder, P., Tiemeier, D. & Enquist, L. *Science* **196,** 175–177.

A position effect in the control of transcription at yeast mating type loci

Kim A. Nasmyth[*], **Kelly Tatchell**[*], **Benjamin D. Hall**[*], **Caroline Astell**[†] **& Michael Smith**[†]

[*] Department of Genetics, SK-50, University of Washington, Seattle, Washington 98195
[†] Department of Biochemistry, Faculty of Medicine, University of British Columbia, 2075 Westbrook Place, Vancouver, British Columbia, Canada V6T 1W5

The two mating type loci MATa *and* MATα *each produce two mRNAs that are transcribed in opposite and diverging directions from central promoters. Silent copies of* MATa (HMRa) *and* MATα (HMLα) *contain identical DNA sequences throughout the transcribed region, yet are not transcribed. It is concluded that sequences to the left of* HMRa (*and probably* HMLα) *must somehow affect transcription initiated at the centre of each locus 700 to 1,400 base pairs away. A possible mechanism for this position effect is discussed.*

THE mating type of a yeast cell, *a* or *α*, is determined by a single locus on chromosome III called *MAT* (ref. 1) which may contain one of two types of DNA sequence[2,3]. In heterothallic cells, *MATa* and *MATα*, are stable mendelian 'alleles' which when hemizygous or homozygous (as in *a* or *α* haploid cells or *a*/*a* or *α*/*α* diploid cells) confer an ability to mate with cells of the opposite mating type and when heterozygous (as in *a*/*α* diploid cells) turn off mating functions and allow the cell to sporulate[1]. wver, in cells containing the homothallism gene (*HO*), *MATa* and *MATα* are interconvertible, switching mating type at almost every cell division until conjugation leads to the formation of a stable *a*/*α* diploid whereupon the *MAT* alleles are stable[4,5].

In addition to the *HO* gene, an *a* to *α* switch requires another locus, *HMLα*, located on the left arm of chromosome III. Similarly, an *α* to *a* interconversion requires *HMRa*, located at the far end of the right arm of chromosome III[6,7]. Genetic[8] and physical[2,3] analyses of the *MAT* loci have revealed that interconversion occurs by a replacement of information present at *MAT* with a copy of information of the opposite mating type present at one of the *HM* loci. In other words, *HMLα* is a silent or unexpressed copy of *MATα*, and *HMRa* is a silent copy of *MATa*. The physical structures and chromosomal location of *MATa* or *MATα*, *HMLα* and *HMRa* are depicted in Fig. 1. At the centre of each locus lies a sequence specific to the mating type, either a 650-base pair *a* specific (Y*a*) or 750-base pair *α* specific (Y*α*) sequence. In most strains *HML* contains Y*α* and *HMR* contains Y*a*. The sequences flanking Y are found at both silent and active loci—a 700-base pair region to the left of Y (called X) and a 230-base pair sequence to its right (called Z1). Two further blocks of DNA are found at both *MAT* and *HML*—700 base pairs adjoining the left end of X (called W) and 90 base pairs adjoining the right end of Z1 (called Z2). W and Z2 are found at *MAT* and *HML* but not at *HMR*, irrespective of whether *a* or *α* information is present.

MAT seems to determine cell type by the control of unlinked *a* or *α* specific genes rather than by direct expression of all structural information required for mating[9]. *MATα* performs this function through the action of at least two genes, *α*1 and *α*2 (ref. 9). The *MATα*1 gene product is necessary for the expres-

0028-0836/81/040244—07$01.00

Table 1 Relative level of *MAT* transcripts in different strains

Genotype		RNA *a*1	*a*2	*α*1	*α*2
HMLα MATa	*HMRa*	++	++	−	−
HMLα MATα	*HMRa*	−	−	++	++
HMLα MATa/MATα	*HMRa*	++	++	−	+
HMLα MATa/matα2	*HMRa*	++	++	++	++
HMLα mata1/MATα	*HMRa*	++	++	++	++
*HMLα MATa SIR*1	*HMRa*	ND	++	±	±
*HMLα MATa SIR*3 or 4 23°C	*HMRa*	ND	++	−	−
*HMLα MATa SIR*3 or 4 34°C	*HMRa*	ND	++	−	+
*HMLα MATα/MATα SAD*1/*SAD*1	*HMRa*	−or±	−or±	++	++
HMLα MATα/a-lethal (HΔ)	*HMRa*	++	++	−	+

The relative levels of each *MAT* transcript in different strains are shown. + + Signifies the highest level seen for a given transcript, + signifies a significant or 3- to 10-fold reduction, ± signifies a 10- to 100-fold reduction and − signifies our inability to detect it. No correspondence is intended for comparisons between transcripts; for these, see the text. Most of the data summarized here can be found by examining Fig. 3. Sometimes ± is written when no such transcript is visible in Fig. 3. This is due to inadequate reproduction of the latter. The Hawthorne deletion removes all DNA on chromosome III between *MAT* and *HMR*, due to an intrachromosomal recombination between these two loci[3]. The *SAD*1 mutation is also due to a recombination between *MAT* and *HMR*, but is an interchromosomal recombination that results in the duplication of all DNA on chromosome III between *MAT* and *HMR* (J. Hicks, personal communication). ND, not determined.

sion of genes unlinked to *MAT* that specify an *α* mating type. *MATα*2, on the other hand, is necessary for repressing *a* specific genes which are otherwise expressed constitutively and would therefore antagonize the expression of an *α* mating type. This model[9] for the determination of mating type implies that *MATa* plays no part in the specification of an *a* mating type as the genes necessary are constitutively expressed in the absence of the *α*2 gene. A gene at *MATa* (*a*1) is, however, necessary for specifying, in conjunction with the *α*2 gene, the non-mating *a*/*α* diploid cell type and therefore the ability to sporulate[10].

The mating type genes, although also present at the silent loci *HMR* and *HML*, are normally exclusively expressed at *MAT*. Recessive mutations at a number of unlinked loci(*SIR*1, 2, 3 and 4)[11,12] allow the silent copies to be expressed, suggesting that *HML* and *HMR* are under negative control of the *SIR* gene products.

Our aim here is to define the process whereby ty genes at *HMRa* and *HMLα* are activated on transposition to *MAT*. We have therefore focused on the location and control of transcripts made from mating type loci. Our results reveal that *MATa* and *MATα* each produce two transcripts (*a*1 and *a*2, *α*1 and *α*2) neither of which is made at their corresponding silent copies *HMRa* and *HMLα*, even though all transcript sequences and their immediate surroundings are present there. More precisely, the sequences flanking the left-hand end of each silent locus is apparently responsible, in conjunction with *SIR* gene products, for preventing the initiation of transcription within the Y regions at the centre of each locus. This situation is reminiscent of position effects in *Drosophila*[13] and similar phenomena recently discovered at the haemoglobin locus in humans[14].

In addition to defining the control of transcription at the silent (*HM*) loci, our results have revealed that *MAT* controls its own transcription; *MATa*1 and *MATα*2 together repress the transcription of *MATα*1 in *a*/*α* diploids, thereby ensuring their non-mating phenotype.

MAT transcripts

We have mapped RNA transcripts in the region of *MAT* by two different methods, R-looping[15] and S_1 nuclease protection[16]. We have used the former to estimate the number and approximate position of transcripts and the latter to discover their direction of transcription and the position of their termini at a nucleotide level.

R-looping

The details of our R-looping are described in Fig. 2 and its legend. Figure 2*A* shows typical examples of R-loops formed between linearized *MATa* or *MATα* plasmid DNA and poly(A)$^+$RNA from *a* or *α* cells. Figure 2*B* shows our conclusions, which can be summarized as follows. First, *a* cells contain two distinct RNA species homologous to *MATa* in roughly equal abundance. One, which corresponds to the *a*1 gene[17], is about 450 base pairs and is transcribed from the Y*a* region with its (centromere proximal) left-hand terminus about 100 base pairs to the right of the X–Y*a* boundary. The other transcript, which we call *a*2, is about 600 base pairs long and is found within the X region with its right-hand (centromere distal) end near the X–Y*a* boundary. Neither *a* transcript seems to have an intervening sequence sufficiently large to be detected by electron microscopy. Second, *α* cells also contain two distinct RNA species homologous to *MATα*. One, which corresponds to the *α*1 gene[17], is about 700 base pairs long and is transcribed from a region extending from the middle of the Y*α* to the end of Z2. The other, which corresponds to the *α*2 gene (ref. 17), is about 720 base pairs long and is transcribed from the X region with its right-hand end just within the Y*α* region. Neither of the *α* transcripts seem to have intervening sequences.

S_1 mapping

Figure 3 shows the strategy used to map the exact 5′ and 3′ termini of all four *MAT* transcripts. We have used the extent of protection from S_1 nuclease, by RNA of single-stranded end-labelled restriction fragments, as a measure of the distance from a labelled restriction site to an RNA terminus. The details of this analysis are described in Fig. 3 and its legend (Table 2 shows the genotype of the strains used in Fig. 3). In the text, therefore, we will confine ourselves to its conclusions.

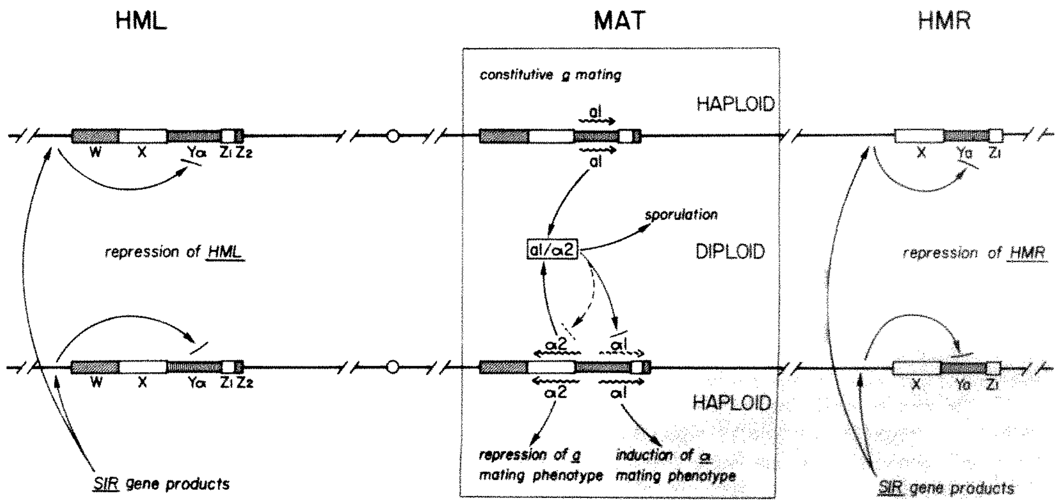


Fig. 1 An overview of yeast mating type loci on chromosome III and the control of cell type by *MATa* and *MATα*. The boxed area refers to the activity of *MATa* and *MATα* in their determination of *a*, *α* or *a*/*α* cell types. The rest of the figure sets these activities in the perspective of other 'silent' copies of *MAT* on the left (*HMLα*) and right (*HMRa*) arms of chromosome III. A description of the comparative structures of *MAT*, *HML* and *HMR* is given in the text. Positive controls are represented by connecting arrows and negative controls by arrows interrupted by a bar.

Table 2 Genotype of strains used

No.	Strain	Genotype		Source
S1	XP8-4A	*MATa ho his6 leu1 met1 trp5-1 gal2 can1*	isogeneic derivatives of XT 1172-S245C	P. Kushner
S2	XP8-10B	*MATα ho his6 leu1 met1 trp5-1 gal2 can1*		P. Kushner
S3	XP11	XP8-4A × XP8-10B		P. Kushner
S4	XR29-10C	*MATa ho sir1-1 ade6 arg4 leu2 trp1 RME*		G. Sprague
S5	2049-1-4a	*MATa ho ste8-59a (sir3) trp1 ade2 leu2*		L. H. Hartwell
S6	2050-5-1a	*MATa ho ste9-62c (sir4) trp1 ade2 leu2*		L. H. Hartwell
S7	265-7-4	*MATa ho hom3 ilv1*		L. H. Hartwell
S8	H293-4-2	*MATa ho trp1*		L. H. Hartwell
S9	K205	H293-4-2 × 9200-15a homozygosed for *trp1* by X rays		This work
S10	K210	K79 × 17-16 homozygosed for *ade2* by X rays		This work
S11	2760	*MATα ho gal12 GAL1 suc ma1 ura1 ade7* *mataΔ ho gal12 gal1 suc ma1 URA1 ADE7*		D. C. Hawthorne
S12	X999-1	*MATα SAD1 ho ade2 his4-12 leu2-27 LYS2 TRP1 ura3 gal2* *MATα SAD1 ho ADE2 his4-12 leu2-27 lys2 trp1 ura3 gal2*		A. Hopper
S13	K79	*MATα ho leu2 trp1*		This work
S14	17-16	*mata1 (a*) ho ade2 ura3 trp1 can1 cyh2*		A. Hopper
S15	G200-15a	*matα2-1 ho sir1-1 ade6 met1 his6 RME*		G. Sprague

(1) The *a*1 and *a*2 genes are transcribed in opposite and diverging directions. The *a*1 transcript is initiated about 90 base pairs to the right of the X–Y*a* (see Figs 3, 4) boundary and extends rightward 510–530 base pairs through the Y*a* region (see Fig. 3, fragment 6). The *a*2 transcript, on the other hand, is initiated about 50 base pairs to the left of the X–Y*a* boundary (Figs 3, 4) and extends leftwards through the X region for 600 base pairs (Fig. 3, fragment 1).

(2) The α1 and α2 genes are also transcribed in opposite and diverging directions. Both are initiated within the Yα region—the α2 gene at a point about 90 base pairs (Figs 3, 4) and the α1 gene at a point about 300 base pairs from the X–Yα boundary (Figs 3, 4). The *a*2 and α2 transcripts have exactly the same termination point (Fig. 3, fragment 1).

(3) All genes, to a greater or lesser extent, have heterogeneous 5′ and 3′ termini. The 5′ end of the *a*2 transcript is a good example of this heterogeneity. Essentially the same spectrum of termini is found, barring a consistent 5-base pair difference, whether S_1 or the single-strand specific exonuclease ExoVII (ref. 16) is used to degrade unprotected fragment 2 DNA (Fig. 3). This suggests that the pattern seen is a true reflection of RNA termini and not an S_1 degradation artefact. We do not know, however, whether this heterogeneity of RNA termini is due to genuine heterogeneous initiation (or termination) or to processing or degradation.

(4) We have mapped all transcript sequences between their termini by the S_1 protection method. This has been achieved either by mapping both 5′ and 3′ termini from a common site within the gene (for example, *MATa*2, Fig. 3) or by also analysing the protection by RNA of end-labelled DNA fragments internal to the gene (such as *MATa*1, Fig. 3). As a result, we have shown that none of the *MAT* transcripts has an intervening sequence between the termini defined by our analysis. This, however, does not exclude splicing of terminal sequences. To exclude the latter in the case of the 5′ end of the *a*2 gene, we have directly compared the results of S_1 and ExoVII digestion and have found no discrepancy except for a consistent 5–7-base pair difference in the length of the protected fragment. ExoVII apparently cannot degrade the last 5 base pairs of a single-strand tail adjacent to duplex DNA (Fig. 3, fragment 2).

Control of *MAT* transcripts

We have studied two aspects of the control of *MAT* transcripts: (1) the regulation of transcription at the 'silent' copies of *MAT*, *HML* and *HMR* and (2) the autoregulation of *MAT* transcription by *MAT* itself.

MATa cells contain neither α1 nor α2 transcripts. Likewise, *MATα* cells contain neither *a*1 nor *a*2 RNAs. This implies that the 'silence' of *HMLα* and *HMRa* is due to a failure in transcription. Recessive mutations at any one of four loci (*SIR* 1, 2, 3 or 4) result in the expression of *HMLα* and *HMRa*[11,12]. Figure 3 (fragment 9) and Table 1 show that, at least in the case of the α2 gene, this is due to a failure to repress *HMLα* transcription. The α2 transcript found in *MATa SIR*1, *SIR*3 or *SIR*4 strains has identical 5′ termini to those found in *a*/α diploids (the most comparable genotype). In the case of *SIR*3 and *SIR*4, we have analysed temperature-sensitive alleles of these loci and have shown that α2 transcription is also temperature sensitive.

SIR gene products, therefore, repress transcription at *HMLα*. It is not clear, however, how this repression is mediated. *MATα* and *HMLα* have identical DNA sequences[18,19] from the left end of W to the right end of Z (a total of 2,500 base pairs), with the apparent promoter region in the middle of this homology. The specific repression of *HMLα* by *SIR* gene products must presumably be effected by flanking sequences outside this homology, beyond the 3′ ends of both α1 and α2 transcripts. The same is presumably true for *HMRa*.

To characterize further the mode of *SIR* gene product action, we have analysed the levels of *a*1 and *a*2 transcription from two variants of *HMRa* which have different flanking sequences; the first is that caused by the Hawthorne deletion[3,20] which replaces the left-hand flanking sequences with those from *MAT* but leaves the right-hand flanking sequences intact, and the second is the converse rearrangement, called *SAD*1 (refs 21, 22), which replaces the right-hand flanking sequences with those from *MAT* but leaves the left-hand flanking sequences intact (J. Hicks, personal communication). Table 1 shows that the former produces the same amount of *a*1 and *a*2 mRNA as *MATa* whereas the latter produces much less of each transcript (<1%). This implies that sequences flanking the left end of *HMRa* are necessary for the *SIR*-mediated repression, whereas those flanking the right end are not. Further explanation of these variant *HMRa* loci is provided in Table 1 legend.

Regulation of *MAT* transcripts by *MATa*1 and *MATα*2

Figure 3 and Table 1 show that the spectrum of *MAT* transcripts in *MATa*/*MATα* diploids is not the sum of its *MATa* and *MATα* parents. Whereas the *a*1 transcript is present in comparable amounts in *a* and *a*/α cells, the α2 transcript is reduced at least fivefold and that of α1 is entirely absent in *a*/α diploids. The α2 transcript starts at three discrete positions in haploids (Fig. 3 and ref. 18), one (<5%) 8 base pairs upstream from the AUG, a second (>90%) 12 base pairs upstream and a third (<5%) 28 base pairs upstream from the AUG. Only the first and second of these transcripts are subject to repression (about 10-fold reduction) in *a*/α diploids. The repression of α1 and α2 transcription in *a*/α diploids requires the action of both *MATa*1 and *MATα*2 genes (Fig. 3, Table 1).

DNA sequence of putative control regions in *MATa* and *MATα*

We have determined the sequence of *MATa* DNA between the starts of *a*1 and *a*2, of *MATα* DNA between the starts of α1 and α2, and the corresponding region of *HML* (see Fig. 4). The following points are noteworthy. First, the region between *a*1 and *a*2 in Y*a* and the corresponding region in Yα have neither extensive sequence homology nor analogous sequence patterns; second, the sequence of the region between α1 and α2 in *MAT*α is identical to that in *HML*α; third, the regions between the structural genes of both *MATa* and *MAT*α contain symmetrical structures (boxed sequences of Fig. 4).

Control of silent copy transcription: a position effect

*MAT*α produces two transcripts that are initiated within 250 base pairs of each other within the unique α region (Yα) and are transcribed in diverging directions. We have shown by DNA sequence analysis[18,19] that for α1 and α2 the entire coding region of the mature transcripts and the sequences between them is virtually identical in *MAT*α and *HML*α. However, *MAT*α and *HML*α differ drastically in their expression of identical *MAT* genes. *HML*α will produce neither α1 nor α2 mature transcript except when mutations at unlinked *SIR* loci are present. The sequence of *MAT*α and *HML*α does diverge beyond the 3′ ends of the transcripts, 800 base pairs to the left of the α2 transcript termination point and at the termination point of the α1 transcript. Somehow, then, sequences downstream of the 3′ ends of the α1 and α2 genes are responsible for the differential repression of transcription at *MAT*α and *HML*α by *SIR* gene products.

A similar situation exists for *MATa* and *HMRa*, although here we have not demonstrated identity at the DNA sequence level. Comparison of the levels of *MATa* transcripts in the Hawthorne deletion[20] and the *SAD*1 mutation[21,22] (Table 1) suggests that sequences flanking the left (centromere proximal) but not the right of *HMRa* are primarily responsible for mediating the repression by *SIR* gene products (Fig. 1).

It is possible that the observed transcripts made from *MATa* and *MAT*α are initiated not, as they seem, within Y*a* and Yα, but from the sequences flanking *MAT*. By this hypothesis, *MAT*α1 would be processed from a long precursor initiated in flanking regions to the left of *MAT*, and *MAT*α2 would be processed from a precursor initiated in flanking regions to its right. One might then interpret the regions of pronounced dyad symmetry in Y*a* and Yα as being for RNA secondary structures involved in processing. The most serious empirical objection to this hypothesis is our observation that *HMRa* with its left-hand flanking sequences replaced by those from *MAT* (the Hawthorne deletion) produces *MAT* levels of both *a*1 and *a*2 transcript, whereas with its right-hand flanking sequences replaced by those from *MAT* (*SAD*1), it barely produces either transcript. In addition to this, the RNA processing hypothesis does not have a convincing explanation for the phenotype of *SIR* mutants. For instance, one would have to postulate that the promoters flanking *MAT* also existed in the nonhomologous sequences flanking *HML*α and *HMRa* but that they were specifically repressed by *SIR* gene products.

If we reject the processing hypothesis, we are forced to conclude that sequences flanking the left end of *HMR* can

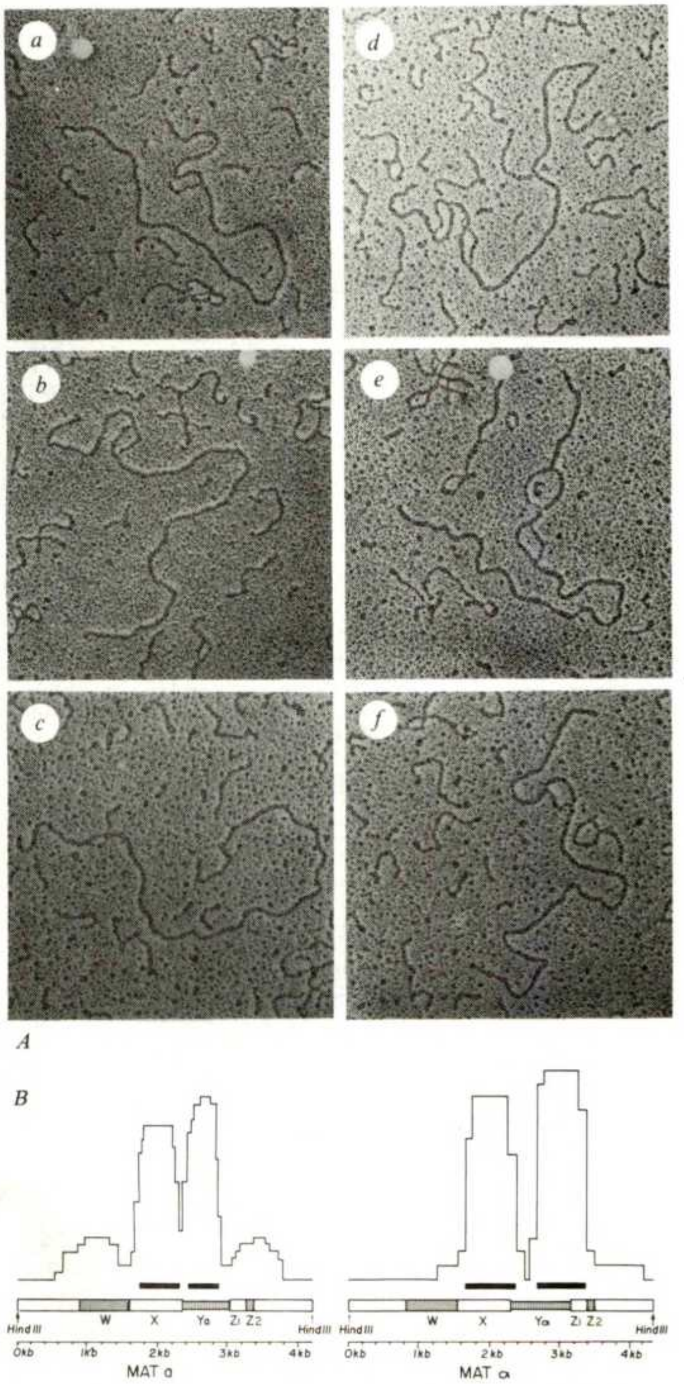


Fig. 2 R-loops formed between linearized *MATa* or α plasmids and poly(A)$^+$ RNA from *a* and α cells. (*a*) *MATa* DNA–*a* RNA: shows *a*1; (*b*) *MATa* DNA–*a* RNA: shows *a*2; (*c*) *MATa* DNA–α RNA: shows α1 (note the tail); (*d*) *MAT*αDNA–α RNA: shows α1 and α2; (*e*) as (*d*); (*f*) *MATa* DNA–α RNA; shows α2 (note the tail). All R-loops were formed between pBR322 containing a 4.3-kilobase *MATa* or *MAT*α *Hin*dIII fragment (inserted into its *Hin*dIII site with the centromere distal side towards the *Bam*Hi site of pBR322) and poly(A)$^+$ RNA made from *a* (S1) or α (S2) cells. The procedure used was essentially that described by Kaback *et al.*[27]. This method involves the use of cross-linked DNA which occasionally results in R-loops being interrupted by a cross-link; see, for instance, (*d*). Plasmid DNA was linearized by restriction with endonuclease *Bam*HI, extracted once with phenol/chloroform, ethanol precipitated, cross-linked with trioxsalen (2–3 cross-links per molecule) and hybridized at 52°C for 15–18 h at $5\ \mu g\ ml^{-1}$ with $1\ mg\ ml^{-1}$ poly(A)$^+$ RNA in a total volume of $25\ \mu l$ sealed in a glass capillary. After hybridization, $4\ \mu l$ of deionized glyoxal (Baker) was added to the pre-cooled (to 0°C) hybridization mixture, which was then incubated at 12–14°C for 2 h. Spreading was perfomed directly without prior purification of RNA loops from the unhybridized RNA, by diluting an aliquot of the glyoxalated mixture with 9 vols of 0°C TE buffer (10 mM Tris-HCl, 1 mM EDTA, *p*H 8.0) and adding $1-2\ \mu l$ of the diluted R-loops to $100\ \mu l$ of spreading solution (50% formamide to be spread on a 20% formamide hypophase). Spreads, their preparation for electron microscopy, their viewing and photography were performed as previously described[2]. A quantitative analysis of the R-loops seen is shown in the lower half of the figure. We have seen R-loops within most regions of the *MAT Hin*dIII fragment, but have only been able to characterize properly transcripts from its central *MAT* region, the others being too rare and having insufficiently precise end points to be characterised by this technique (see later S_1 mapping). *MATa* and *MAT*α both produce two discrete transcripts that are apparent in the histograms beneath the R-loop pictures. The apparent difference between *MATa* and *MAT*α in the occurrence of flanking transcripts is probably due to the *MAT* specific transcripts from the former being at a lower level (of detection, possibly) than those from the latter. Underneath the histogram is a map of *MAT* with the position of *MAT* transcripts placed on it. These positions were estimated by averaging the end points of the molecules observed. The *MAT*α transcripts both overlap homologous (X or Z) and unique (Yα) regions. Thus, one might predict the occurrence of different types of R-loop if α RNA was hybridized instead to *MATa* DNA. One should find in the case of α2 a slightly shorter R-loop with a short RNA tail and in the case of α1 a very much shorter R-loop with a long RNA tail. *We have consistently found the latter structure* in the predicted position but have only found individual examples of the former, (see (*c*) and (*f*), respectively).

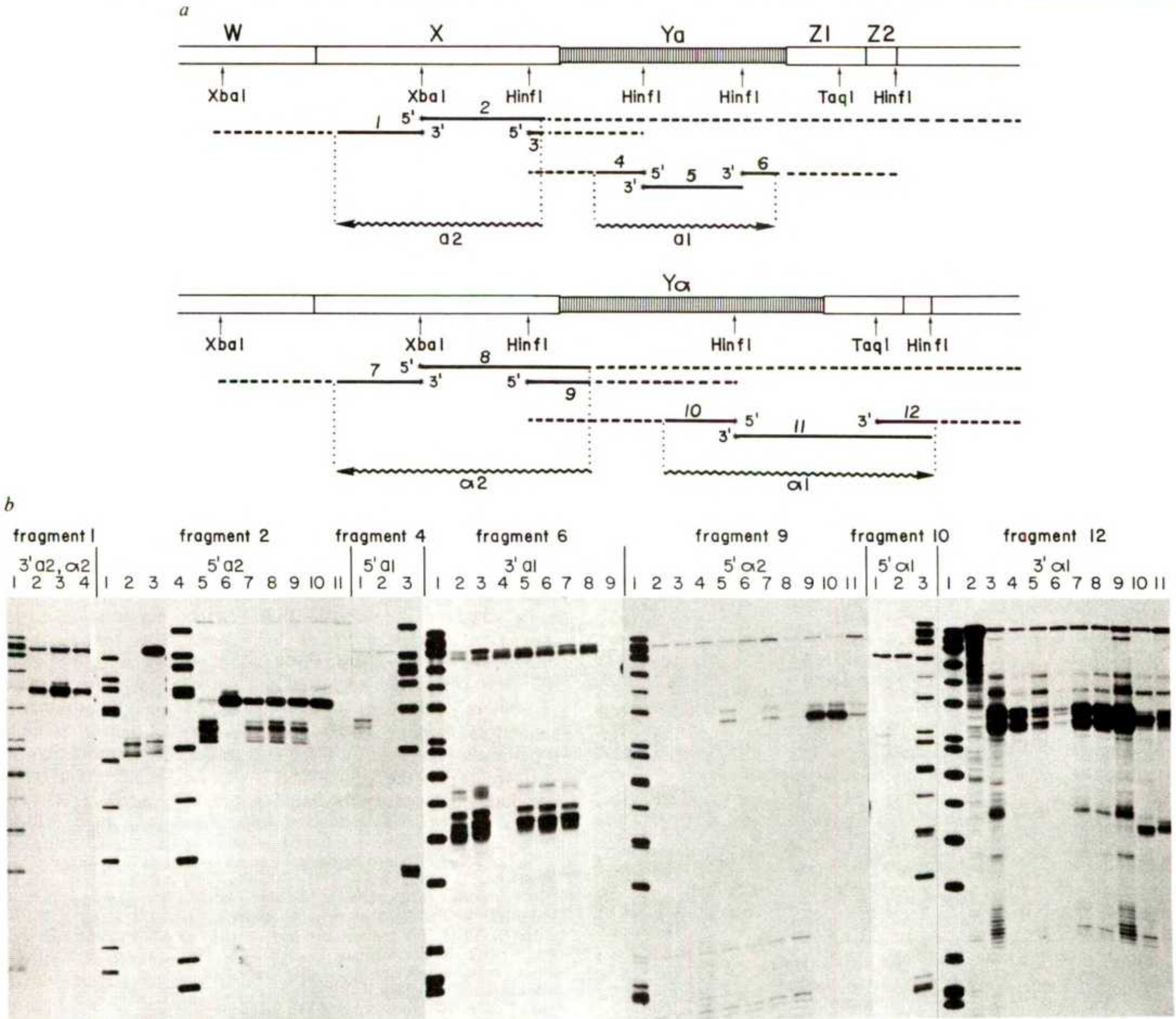


Fig. 3 S_1 transcription mapping. The upper part of the figure shows the strategy by which we have mapped the exact 5′ and 3′ termini of all four *MAT* transcripts. By hybridizing total poly(A) RNA to end-labelled separated strands of restriction fragments, then measuring the lengths of the ^{32}P-DNA strands remaining after nuclease digestion, we have measured the distance from each RNA terminus to a physically mapped restriction site. The single-stranded, end-labelled fragments used as probes are given numbers (1–12). The *Xba*I, *Hin*fI and *Taq*I fragments were derived from wild-type *MATa* or *MATα* DNA (*Hin*dIII fragments) cloned in pBR322. Fragment 2 was derived from an *in vitro* manufactured mutant with a *Xho*I site within Y*a* (*mata*23, ref. 17). 5′* at the end of each fragment indicates an end labelled by polynucleotide kinase and 3′* indicates an end labelled with Klenow DNA polymerase. Solid lines indicate stretches of DNA protected from S_1 nuclease by RNA whereas dashed lines indicate sensitivity. Apart from the four major transcripts presented here, we have evidence for at least three more transcripts within the *MAT* region (W–Z of fig. 2). One of these, made at least in *a* cells (see R-looping), covers part of the W region. Another (called *MAT*3), made from both *MATa* and *MATα* in *a*/*α* cells, starts some point before the *Taq*I site of Z1 and ends at the same points as the *MATα*1 transcript. S_1 protection experiments comparing *a*/*α* cells with *a* lethal (Hawthorne deletion)/*α* cells[20] show that half of these *MAT*3 transcripts cease to have any homology to *MATa* from a point 88 base pairs from the *Taq*I site. This implies that the Z1–Z2 boundary occurs at this point. A third minor transcript, made at least in *α* cells, starts somewhere to the right of the *Taq*I site to the right of the Z1 *Taq*1 site, and ends somewhere near or close to the left of the latter. We do not know the biological significance of these minor transcripts. Only those restriction sites used for S_1 mapping are shown. The lower half of the figure presents the autoradiograms of end-labelled fragments. The markers are either ΦX174 or pBR322 cut with *Hin*fI and end-labelled with Klenow DNA polymerase. The numbers in parentheses after the *MAT* genotypes refer to the strain number of Table 2. Unless otherwise stated, all digestions were performed with S_1 nuclease and cells were grown at 30°C in YEPD (1). **Fragment 1 (or 7):** The 3′ end of *MATa*2 and *MATα*2. Protection of fragment 1 or 7 by RNA at 1 mg ml^{-1}. (1) ΦX174, (2) *MATa* (S1), (3) *MATα* (S2), (4) *MATa*/*MATα* (S3). **Fragment 2:** The 5′ end of *MATa*2. Protection of fragment 2 by RNA at 1 mg ml^{-1} (lanes 2, 3) or 2.5 mg ml^{-1} (lanes 5–10). (1) ΦX174; (2) *MATa* (S1)—S_1 nuclease; (3) *MATa* (S1)—ExoVII; (4) ΦX174; (5) *MATa* (S1); (6) *MATα* (S2); (7) *MATa*/*MATα* (S3); (8) *mata*1/*MATα* (S10); (9) *MATa*/*matα*2 (S9); (10) *MATα*/*MATα SAD*1/*SAD*1 (S12); (11) no RNA. The intense band(s) at 400 nucleotides in lanes 6–10 corresponds to *α*2 RNA, which is also homologous to the probe up to the X–Y boundary. **Fragment 4:** The 5′ end of *MATa*1. Protection by RNA at 1 mg ml^{-1}. (1) *MATa* (S1); (2) *MATα* (S2); (3) ΦX174. **Fragment 6:** The 3′ end of *MATa*1. Protection by RNA at 2.5 mg ml^{-1}. (1) ΦX174; (2) *MATa* (S1); (3) *MATa* (S7); (4) *MATα* (S2); (5) *MATa*/*MATα* (S3); (6) *mata*1/*MATα* (S10); (7) *MATa*/*matα*2 (S9); (8) *MATα*/*MATα SAD*1/*SAD*1 (S12); (9) no RNA. **Fragment 9:** The 5′ end of *MATα*2. Protection by RNA at 1 mg ml^{-1}. (1) ΦX174; (2) *MATa* (S1) grown at 23°C; (3) *MATa* (S1) 34°C; (4) *MATa SIR*3 (S5) 23°C; (5) *MATa SIR*3 (S5) 34°C; (6) *MATa SIR*4 (S6) 23°C; (7) *MATa SIR*4 (S6) 34°C; (8) *MATa SIR*1-1 (S4) 23°C; (9) *MATα* (S2) 23°C; (10) *MATα* (S2) 34°C; (11) *MATa*/*MATα* (S3) 30°C. **Fragment 10:** The 5′ end at *MATα*1. Protection by RNA at 1 mg ml^{-1}. (1) *MATα* (S2); (2) *MATa*/*MATα* (S3); (3) pBR322. **Fragment 12:** The 3′ end of *MATα*1. All lanes except lane 2 show protection of fragment 12 by RNA at 2.5 mg ml^{-1}. Lane 2 shows protection of the opposite strand to fragment 12. (1) ΦX174; (2) *MATα* (S2); (3) *MATα* (S2); (4) *MATα* (S2) 0.5 mg ml^{-1} RNA; (5) *MATa* (S1); (6) *MATa*/*MATα* (S3); (7) *mata*1/*MATα* (S10); (8) *MATa*/*matα*2 (S9); (9) *MATα*/*MATα SAD*1/*SAD*1 (S12); (10) *MATa sir*1-1 (S4); (11) *MATa sir*3 (S5) grown at 34°C. Lanes 5 and 6 show bands characteristic of *α*1 termini even though no transcripts with the 5′ sequences of *α*1 have been detected in these strains (see fragment 10). These termini probably belong to a transcript (*MAT*3?) that starts in Z and terminates at the same points as *α*1 (see below). Klenow DNA polymerase and T4 polynucleotide kinase were used to label the 3′ and 5′ ends, respectively, of restriction fragments which were then strand separated by electrophoresis. Kinase reactions and strand separations were done as described by Maxam and Gilbert[28] and the Klenow polymerase I reaction and all further procedures as described by K.A.N. *et al.*[18]. Poly(A)$^+$ RNA and single-strand end-labelled DNA fragments were annealed at 65°C followed by S_1 (or ExoVII) digestion. After digestion the samples were ethanol precipitated, redissolved in NaOH–urea loading buffer and run on 6.0% or 8.0% acrylamide sequencing gels[28]. Gels were exposed to X-ray film plus intensifying screen at −70°C for 1–14 days.

somehow affect the initiation of transcription within Y*a*, a situation reminiscent of position effects in *Drosophila*[13] and similar phenomena at the haemoglobin locus in humans[14].

Possible mechanisms for the position effect

Hypothetical mechanisms for the position effect can be divided into two distinct types: one in which the DNA sequence is solely responsible for the effect and another in which chromosome structure is also involved.

The first type of mechanism could, for example, involve the action of an enzyme like the bacterial type I restriction endonucleases[23], which can recognize a particular sequence on a DNA molecule and then modify or cut the same molecule at a point distant from the recognition site. It is possible that the recognition site for such an enzyme (a *SIR* gene product?) exists within the left-hand flanking region of *HMLα* and *HMRa* and that DNA modification (or some other interaction) subsequently occurs within Y*a* or Y*α*.

There are, of course, many variants of the second type of mechanism in which chromosome structure has an important role. These range from the very general, that *HML* and *HMR* sit within a heterochromatic region, to the very specific, that the precise composition or phasing of nucleosomes is responsible. One can, in fact, distinguish two general pathways whereby the left-hand flanking sequences of *HML* and *HMR* could influence the structure of Y*a* and Y*α*. Information from the flanking sequences could be transmitted to Y*a* or Y*α* either directly by some form of physical bridge or indirectly along the axis of the DNA molecule separating them. Higher order chromosome folding might be involved in the case of the former, whereas the precise phasing of nucleosomes might be involved in the latter.

We propose that of all the possible explanations for the position effect, the most readily tested is the notion that the sequences flanking the left end of *HML* and *HMR* are responsible, possibly in conjunction with *SIR* gene products, for initiating, from left to right, a precise phasing of nucleosomes such that promoters of transcription at Y*a* or Y*α* are either hidden or sites for repressors of transcription (*SIR* gene products) are exposed at these two loci but not at *MAT*. *MAT*, having different flanking sequences, might either be phased differently or not at all (as might be the case for a repressor model). The phasing of nucleosomes could affect the availability or activity of control sequences at Y*a* or Y*α* at two levels. At a gross level, it could be responsible for bringing sequences into juxtaposition or not; for instance, by their being wrapped around the same rather than a neighbouring core particle. At a finer level, it could also, or instead, be responsible for the exact orientation of sequences with respect to their coiling around a core particle; for instance, whether they face in or out.

Possibly relevant to the phasing hypothesis is the recent observation by Hicks (personal communication) that restriction fragments containing *HMLα* and *HMRa* both contain replication origins whereas that containing *MAT* does not. If the inheritance of nucleosomes is conservative, one must postulate some means by which the naked daughter chromatid is rapidly covered with in-phase nucleosomes. This could be more readily achieved if the replication origins are also orgins of phasing, as found in the case of SV40[25]. The proximity of replication origins at the silent copies may be sufficient to effect the correct phasing. On the other hand, one or more *SIR* gene products may be involved in initiating as well as maintaining phasing.

Whatever the mechanism of the position effect on the expression of *MAT* genes, it is worth pointing out possible parallels with various phenomena associated with mating type interconversion. For instance, how does a homothallic yeast cell distinguish *MAT* from *HML* as the target of an interconversion when they are so extensively homologous? It is tempting to speculate that the same property that distinguishes their transcription may also be used to distinguish their roles as donor and recipient in an interconversion of mating type. This notion raises the question of the role of recombinant mating type loci (such as the *MAT*/*HMR* recombinant caused by the Hawthorne deletion) in an interconversion; that is, whether they serve as donors or recipients. Pursuing the analogy still further, a conservative inheritance of nucleosomes at *MAT* or its silent copies may be responsible for the ability of mother cells, but not daughter cells, to undergo mating type interconversion[4,5].

Regulation of *MAT* transcription and the determination of cell type in yeast

Genetic analyses have identified three genes involved in the specification of cell type in yeast: *α*1, *α*2 and *a*1 (refs 9, 10). The present investigation has shown that *MATa* and *MATα* each produce two major transcripts starting in or near their Y regions, one going leftwards and another rightwards. A series of experiments in which plasmids containing *MATa* or *MATα* were mutated *in vitro* and subsequently tested for *in vivo* function have shown that the left *α* transcript corresponds to the *α*2 gene, the right *α* transcript to the *α*1 gene, and the right *a* transcript to the *a*1 gene[17]. This leaves the left (*a*2) transcript unassigned to a known complementation group. The same series of experiments implied that this transcript was, however, not essential for any of the known functions performed by plasmids carrying a wild-type *MATa* fragment; that is, deletions within the *a*2 region had no effect on the plasmid's ability to suppress *α* mating functions or to induce sporulation (in *a*/*α* cells). We therefore believe the *α*1, *α*2 and *a*1 genes to be both necessary and sufficient for specifying all three yeast cell types: the *a* mater, the *α* mater and the non-mating but sporulating *a*/*α* cell[1].

How, then, do these three genes determine cell type? Strathern *et al.*[9] have suggested that the *α*1 gene is necessary for promoting the expression of unlinked *α* specific mating genes and the *α*2 gene for repressing the expression of unlinked and constitutive *a* specific mating genes. With both of these active, as in *α* haploids, *α* specific mating genes will be expressed but *a* specific mating genes not. In the absence of either, as in *a* haploids, only the constitutive *a* specific mating genes will be expressed. The presence, albeit lower, of *α*2 and complete absence of *α*1 transcripts in *a*/*α* diploids therefore accounts for the non-mating phenotype of these cells (Fig. 1).

The repression of *MATα* transcription that results in the complete absence of *α*1 and lowering of *α*2 transcripts in *a*/*α* cells requires both the *a*1 and *α*2 genes. The sequence between the promoters of the *α*1 and *α*2 genes contains a symmetrical region that may be the target of the hypothetical repressor of *α*1 and *α*2. No similar structure is seen near the start of the *MATa*1 transcript, which we have shown to be refractory to the repression by *a*1 and *α*2,

The *a*1 and *α*2 genes are necessary not only for suppressing mating functions in *a*/*α* diploids but also for inducing their ability to sporulate. This function cannot be just a consequence of turning off *α*1 and turning down *α*2 because it can be expressed in circumstances in which the repression of *α*1 and *α*2 is absent[21]. It is, therefore, an independent property of the *a*1/*α*2 state (Fig. 1). Our analysis of transcription in *a*/*α* *SAD*1/*SAD*1 strains suggests that the induction of sporulation by *a*1/*α*2 requires less *a*1 gene product than does the repression of *α*1 and *α*2 transcription. The existence of recessive mutations (for example, *CSP*, ref. 26) that allow sporulation of *a*/*a* or *α*/*α* cells suggests that sporulation is normally under negative control. It is tempting to speculate that the *a*1/*α*2 repression of *α*1 transcription and induction of sporulation are mediated by a common mechanism; that is, the transcription of *CSP* is repressed by *a*1/*α*2 in a manner similar, though not identical, to that of *α*1.

As the amino acid sequence of all *MAT* gene products has recently been deduced from their DNA sequence[18,19], we may speculate on the physical mechanisms by which they control cell type. We know that some of the consequences of their action are regulation of transcription. However, we do not yet know whether they act so themselves or through intermediates. With this in mind, it is interesting that the *MAT* gene products are all basic proteins that may be capable of binding to DNA.

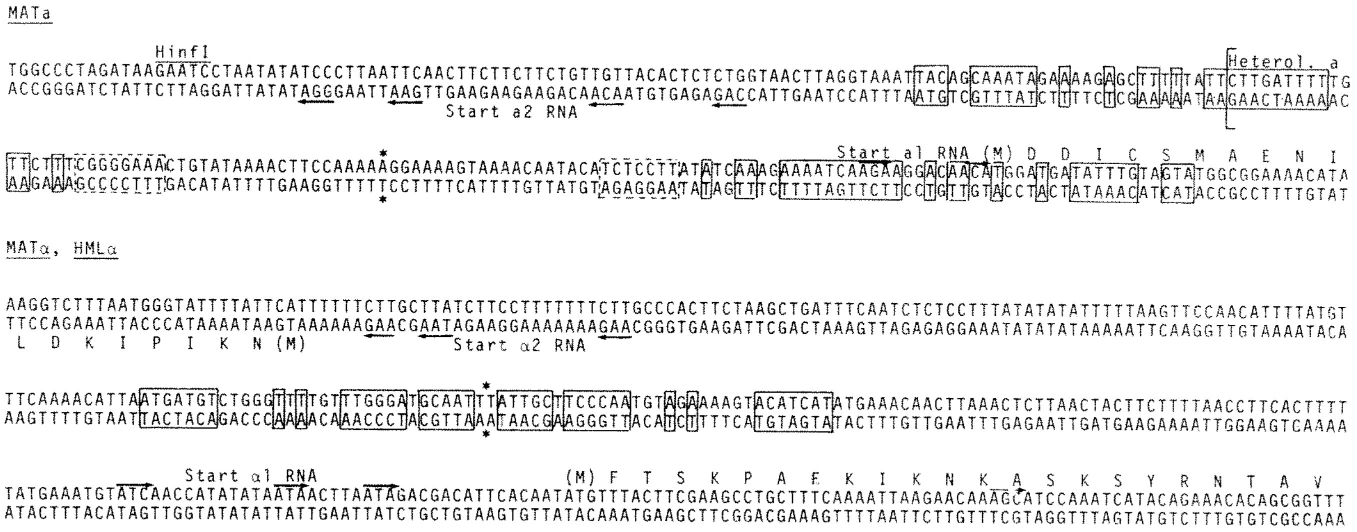

MATa
HinfI
TGGCCCTAGATAAGAATCCTAATATATCCCTTAATTCAACTTCTTCTTCTGTTGTTACACTCTCTGGTAACTTAGGTAAATTACAGCAAATAGAAAAGAGCTTTTTATTCTTGATTTTTG
ACCGGGATCTATTCTTAGGATTATATAGGGAATTAAGTTGAAGAAGAAGACAACAATGTGAGAGACCATTGAATCCATTTAATGTCGTTTATCTTTTCTCGAAAAATAAGAACTAAAAAC
Start a2 RNA
Heterol. a
Start a1 RNA (M) D D I C S M A E N I
TTCTTTCGGGGAAACTGTATAAAACTTCCAAAAAGGAAAAGTAAAACAATACATCTCCTTATATCAAAGAAAATCAAGAAGGACAACATGGATGATATTTGTAGTATGGCGGAAAACATA
AAGAAAGCCCCTTTGACATATTTTGAAGGTTTTTCCTTTTCATTTTGTTATGTAGAGGAATATAGTTTCTTTTAGTTCTTCCTGTTGTACCTACTATAAACATCATACCGCCTTTTGTAT
MATα, HMLα
AAGGTCTTTAATGGGTATTTTATTCATTTTTTCTTGCTTATCTTCCTTTTTTTCTTGCCCACTTCTAAGCTGATTTCAATCTCTCCTTTATATATATTTTTAAGTTCCAACATTTTATGT
TTCCAGAAATTACCCATAAAATAAGTAAAAAAGAACGAATAGAAGGAAAAAAAGAACGGGTGAAGATTCGACTAAAGTTAGAGAGGAAATATATATAAAAATTCAAGGTTGTAAAATACA
L D K I P I K N (M)
Start α2 RNA
TTCAAAACATTAATGATGTCTGGGTTTTGTTTGGGATGCAATTTATTGCTTCCCAATGTAGAAAAGTACATCATATGAAACAACTTAAACTCTTAACTACTTCTTTTAACCTTCACTTTT
AAGTTTTGTAATTACTACAGACCCAAAACAAACCCTACGTTAAATAACGAAGGGTTACATCTTTTCATGTAGTATACTTTGTTGAATTTGAGAATTGATGAAGAAAATTGGAAGTCAAAA
Start α1 RNA
(M) F T S K P A F K I K N K A S K S Y R N T A V
TATGAAATGTATCAACCATATATAATAACTTAATAGACGACATTCACAATATGTTTACTTCGAAGCCTGCTTTCAAAATTAAGAACAAAGCATCCAAATCATACAGAAACACAGCGGTTT
ATACTTTACATAGTTGGTATATATTATTGAATTATCTGCTGTAAGTGTTATACAAATGAAGCTTCGGACGAAAGTTTTAATTCTTGTTTCGTAGGTTTAGTATGTCTTTGTGTCGCCAAA

Fig. 4 DNA sequences around the starts of *MAT* transcripts. The DNA sequences around and between the starts of *MATa* and *MATα* transcripts are shown. The sequences were determined by the method of Maxam and Gilbert[28]. The sequencing strategy will be described elsewhere along with a more complete sequence of the region[19]. In *MATa*, the X–Y*a* boundary is marked by a vertical bracket. The equivalent boundary for *MATα* and *HMLα* (X–Y*α*) is 48 base pairs beyond the left end of the sequence shown. Putative coding sequences are signified using the one-letter amino acid code. Those of the *a*2 transcript are not marked because there is no evidence that this transcript encodes a functional protein. Arrows mark the start points of RNAs as determined by S_1 mapping. Each transcript has two or more start points. The boxed sequences are regions that have a dyad symmetry around an axis marked with an asterisk. At each end of the *α* symmetrical region is a 7-base pair inverted repeat sequence that would extend the region of perfect symmetry except for 7 extra base pairs inserted between it and the left-hand end of the symmetrical region. The dashed boxes in the *MATa* sequence are regions of symmetry with respect to purine and pyrimidine content.

We thank Hal Weintrub for helpful discussions on nucleosome phasing, Don Hawthorne and George Sprague for supplying crucial strains, and Jane Garnett for help in compiling the paper. This investigation was supported by a grant to B.D.H. from the National Institute of Medical Sciences, a Jane Coffin Childs postdoctoral fellowship awarded to K.A.N. and a Damon Runyon-Walter Winchell Cancer fund postdoctoral fellowship to K.T.

Received 5 August; accepted 17 October 1980.

1. Mortimer, R. K. & Hawthorne, D. C. in *The Yeasts* Vol. 1 (eds Rose, A.H. & Harrison, J. S.) 385–460 (Academic, New York 1969).
2. Nasmyth, K. A. & Tatchell, K. *Cell* **19**, 753–764 (1980).
3. Strathern, J. N., Spatola, E., McGill, C. & Hicks, J. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2839–2843 (1980).
4. Hicks, J. B. & Herskowitz, I. *Genetics* **83**, 245–258 (1976).
5. Strathern, J. N. & Herskowitz, I. *Cell* **17**, 371–381 (1979).
6. Naumov, G. I. & Tolstorukov, I. I. *Genetika* **9**, 82–91 (1973).
7. Harashima, S., Nogi, Y. & Oshima, Y. *Genetics* **77**, 639–650 (1974).
8. Hicks, J. B., Strathern, J. N. & Herskowitz, I. in *DNA Insertion Elements Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 457–462 (Cold Spring Harbor Laboratory, New York, 1977).
9. Strathern, J. N., Hicks, J. & Herskowitz, I. *J. molec. Biol.* (in the press).
10. Kassir, Y. & Simchen, G. *Genetics* **82**, 187–202 (1976).
11. Rine, J., Strathern, J., Hicks, J. & Herskowitz, I. *Genetics* **93**, 877–901 (1979).
12. Klar, A. J. S., Fogel, S. & Macleod, K. *Genetics* **93**, 37–50 (1979).
13. Lewis, E. B. *Adv. Genet.* **3**, 73–113 (1950).
14. Van der Ploeg, L. H. T. *et al. Nature* **283**, 637–642 (1980).
15. Thomas, M., White, R. L. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2294–2298 (1976).
16. Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1274–1278 (1978).
17. Tatchell, K., Nasmyth, K. A., Hall, B. D., Astell, C. & Smith, M. *Cell.* (submitted).
18. Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. & Smith, M. *Cold Spring Harb. Symp. quant. Biol.* **45** (in the press).
19. Astell, C. *et al. Cell.* (submitted).
20. Hawthorne, D. C. *Genetics* **48**, 1727–1729 (1963).
21. Hopper, A. K. & Mackay, V. *Molec. gen. Genet.* (in the press).
22. Kassir, Y. & Herskowitz, I. *Molec. gen. Genet.* (in the press).
23. Roberts, R. J. *CRC Crit. Rev. Biochem.* **4**, 123–164 (1976).
24. Leffak, I. M., Grainger, R. & Weintrub, H. *Cell* **12**, 837–845 (1977).
25. Saragosti, S., Moyne, G. & Yaniv, M. *Cell* **20**, 65–73 (1980).
26. Hopper, A. K. & Hall, B. D. *Genetics* **80**, 41–59 (1975).
27. Kaback, D. B., Angerer, L. M. & Davidson, N. *Nucleic Acids Res.* **6**, 2499–2517 (1979).
28. Maxam, A. & Gilbert W. *Meth. Enzym.* **65**, 499–560 (1980).

Homology and non-homology at the yeast mating type locus

George F. Sprague Jr, Jasper Rine* & Ira Herskowitz

Institute of Molecular Biology and Department of Biology, University of Oregon, Eugene, Oregon 97403

Four mutations of the α mating type locus of Saccharomyces cerevisiae *have been analysed to determine their relationship to the* a *mating type locus.* Matα$^+$ *recombinants are produced by* matα2$^-$/MATa *but not by* matα1$^-$/MATa *diploids.* MATα *and* MATa *thus contain regions of homology (coding for at least part of* MATα2*) and regions of non-homology (coding for at least part of* MATα1*)—the genetic determinant for cell type is larger than the non-homologous sequence seen by DNA–DNA heteroduplexes and genetic analysis. The segment transposed in mating type interconversion includes both types of sequence.*

THE mating type locus determines which of three distinct and complex cell types (*a*, *α* and *a*/*α*) is expressed by the yeast, *Saccharomyces cerevisiae*. It has been proposed that alleles of this locus, *MATa* and *MATα*, code for different regulators controlling the expression of unlinked genes necessary for mating and sporulation[1–3]. Several lines of evidence suggest that *MATa* and *MATα* are 'complex' alleles, that is, are genetic alternatives coding for several different functions. First, there are at least two complementation groups in *MAT*α, α1 and α2; only one complementation group is known for *MATa*. *MAT*α1 and *MAT*α2 are required for mating in haploids; *MAT*α2 and *MATa* are required for sporulation in diploids[2-5]. Second, heteroduplex analysis of cloned *MATa* and *MAT*α DNA (as well as cloned cryptic mating type information from the *HML*α and *HMRa* loci) reveals a region of non-homology (region Y) between *a* and α information of ~750 base pairs (Y*a* for *a* and Yα for α). Heteroduplexes between *HML*α and *MATa* and between *HMRa* and *MAT*α show that the non-homology region is flanked by regions of apparent homology of at least 700 base pairs on one side (the X region) and at least 200 base pairs on the other side (the Z region)[6,7]. The physical picture of the mating type locus derived from heteroduplex analysis does not provide information on the coding regions within the mating type locus. Thus, they could be entirely within the non-homology (Y), entirely within the regions of apparent homology (X and Z) or distributed throughout X, Y and Z regions. We have made detailed studies of four mutations of the α mating type locus, two in each of the known complementation groups, to determine their genetic relationship to the *a* mating type locus. To investigate whether *mat*α mutations are in regions of homology or non-homology with respect to *MATa*, we have constructed a series of *mat*α$^-$/*MATa* diploids and have analysed the meiotic products for *MAT*α$^+$ recombinants. If a mutation is in a region of *MAT*α that is non-homologous with *MATa*, genetic rescue of the wild-type allele from *MATa* should not be possible, but genetic rescue should be possible if the mutation is in a region that is homologous with *MATa*. The mutations analysed here (isolated by MacKay and Manney[8]) all differ: *mat*α1-5 causes a temperature-sensitive mating phenotype (G. F. S. and J. R., unpublished), *mat*α1-2 is an *amber* mutation[9], and *mat*α2-1 and *mat*α2-4 have different effects on sporulation (G. Sprague and J. Strathern, unpublished). We report here results obtained with three types of diploid: *mat*α1$^-$/*MATa*, *mat*α2$^-$/*MATa*, and *mat*α1$^-$/*mat*α2$^-$, which show that the wild-type alleles of *mat*α2 mutations but not of *mat*α1 mutations are rescued from *MATa*.

---

* Present address: Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305.

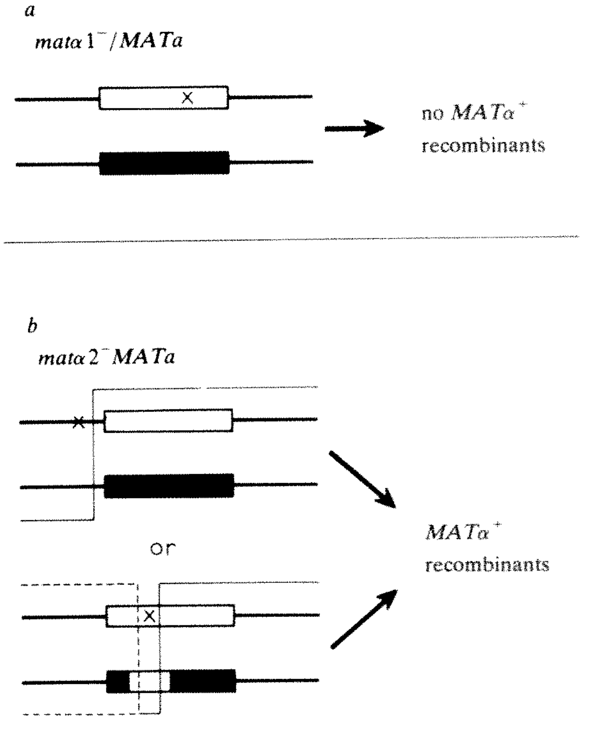


Fig. 1 Interpretation of genetic rescue experiments between *mat*α mutations and *MATa*. X, sites of *mat*α1 or MATα mutations. *a*, Failure to obtain *MAT*α$^+$ recombinants in crosses between *mat*α1-5 and *MATa*. Filled and open rectangles indicate non-homologous segments of *MATa* and *MAT*α, respectively. *b*, Production of *MAT*α$^+$ recombinants in crosses between MATα$^-$ and *MATa*. Upper diagram: symbols as in *a*. Lower diagram: a region of homology is located between sequences that are non-homologous in *MATa* and *MAT*α.

Production of *MAT*α$^+$ segregants from *mat*α$^-$/*MATa* diploids

Segregants that mate as α cells are found at the same low frequency (~5×10^{-5}) among the meiotic products of two *mat*α1$^-$/*MATa* diploids carrying different *mat*α1 mutations—*mat*α1-5 and *mat*α1-2 (strains DG1 and DG2, respectively, Table 1). Subsequent crosses show that the α mating phenotype is due to the presence of a wild-type α mating type locus in these recombinants (see Table 1 legend). Because these *MAT*α$^+$ segregants may have arisen by recombination between the *mat*α1 allele and cryptic α information at the *HML*α locus rather than by recombination between *mat*α1 and *MATa*, we have constructed a *mat*α1-5/*MATa* diploid lacking cryptic α information. Such a strain (DG3, *mat*α1/*MATa* *HMLa*/*HMLa* *HMRa*/*HMRa*) does not yield *MAT*α$^+$ segregants. This result indicates that the wild-type allele of the *mat*α1-5 mutation cannot be rescued from *MATa*.

In contrast to these results, *MAT*α$^+$ segregants arise at high frequency (~2×10^{-3}) from *mat*α2$^-$/*MATa* diploids carrying either of two different *mat*α2 mutations, *mat*α2-4 and *mat*α 2-1. This result has been documented most thoroughly for the *mat*α2-4 mutation. In particular, the production of *MAT*α$^+$ segregants from *mat*α2-4/*MATa* diploids is not dependent on the presence of the *HML*α allele (Table 1: compare strain DG5, which is homozygous for *HMLa* and *HMRa*, with strain DG4, which is *HML*α *HMRa*). Two observations indicate further that the *MAT*α$^+$ segregants are the products of recombination between *mat*α2-4 and *MATa*. First, the frequency of *MAT*α$^+$ segregants is at least 100-fold higher than the reversion frequency of the *mat*α2-4 mutation (HG2, Table 1). Moreover, half of the *MAT*α$^+$ segregants are recombinant at the closely linked *CRY*1 locus, which indicates that *MAT*α$^+$ has been generated by a recombination event rather than by reversion.

It is possible that the observed frequency of *MAT*α$^+$ segregants from *mat*α2-4/*MATa* diploids is an overestimation of the true frequency, because DG4 and DG5 are defective in the *MAT*α2 function and therefore sporulate poorly. Wild-type *MAT*α recombinants or revertants that arise in DG4 or DG5 may be enriched among the spore progeny due to efficient *MAT*α$^+$-promoted sporulation. To test this possibility, we determined the frequency of *MAT*α$^+$ segregants from a *mat*α2-4/*MATa* diploid (DG6) homozygous for a mutation, *sir*1-1, which permits *in situ* expression of the cryptic mating type information[10]. This diploid sporulates well because wild-type *MAT*α2 function is provided by α information at *HML*α. *MAT*α$^+$ segregants arise from DG6 at the same frequency as from DG4 and DG5 (Table 1). A *mat*α2-1/*MATa* *sir*1-1/*sir*1-1 diploid likewise yields *MAT*α$^+$ recombinants at high frequency (DG7, Table 1). Consequently, the frequency of *MAT*α$^+$ segregants from diploids DG4, 5, 6 and 7 seems to be a true reflection of recombination between *MATa* and the MATα$^-$ mutations and is not biased by a selection for revertants or recombinants.

To evaluate the significance of the recombination frequencies reported above, we have determined the frequency of recombination between two *mat*α mutations. The frequency of recombination in *mat*α1$^-$/*mat*α2$^-$ diploids (~10^{-2}, DG8 and DG9, Table 1) is similar to that obtained from MATα$^-$/MATa diploids. This value for recombination across homologous mating type loci places an upper limit on the frequency of

*MATα*2$^+$ recombinants expected from *matα*2-1/*MATa* or *matα*2-4/*MATa* diploids (for example, DG4, 5, 6, 7 above), in which *MATα*2$^+$ recombinants are formed by a recombination event occurring between the *matα* mutation and the end of the *MATa* non-homology region.

Table 1 Frequency of α mating type clones obtained

Strain	Genotype	Frequency of α-maters Clone	 Frequency
DG1	*matα*1-5 / *MATa*	A B	$\sim 4\times10^{-5}$ (2) $\sim 6\times10^{-5}$ (3)
DG2	*matα*1-2 / *MATa*	A B	$\sim 1\times10^{-4}$ (2) $\sim 4\times10^{-5}$ (1)
DG3	*HMLa matα*1-5 *HMRa* / *HMLa MATa HMRa*	A B	$<1.5\times10^{-5}$ (0) $<1.2\times10^{-5}$ (0)
DG4	*matα*2-4 / *MATa*	A B	1.6×10^{-3} (54) 2×10^{-3} (60)
DG5	*HMLa matα*2-4 *HMRa* / *HMLa MATa HMRa*	A B	1.3×10^{-3} (21) 1.3×10^{-3} (48)
DG6	*matα*2-4 *sir*1-1 / *MATa sir*1-1	A B	3×10^{-3} (54) 4×10^{-3} (54)
DG7	*matα*2-1 *sir*1-1 / *MATa sir*1-1	A B	3×10^{-3} (21) 2×10^{-3} (27)
DG8	*matα*1-5 / *matα*2-4		$\sim 4\times10^{-3}$ (2)
DG9	*matα*1-5 / *matα*2-1		9×10^{-3} (9)
HG1	*matα*1-5	A B	$<5\times10^{-5}$ (0) $<3\times10^{-5}$ (0)
HG2	*HMLa matα*2-4 *HMRa*	A B	$<4\times10^{-5}$ (0) $<4\times10^{-5}$ (0)

Strains and markers: Diploid strains were constructed by prototroph selection. Genotypes were confirmed by analysis of the segregants. HG1 and HG2 were the *matα* parents of DG1 and DG5. All diploids except DG1 were heterozygous at the *CAN*1 locus, which segregates independently from *MAT* (ref. 8). DG1, 2, 3 and 5 were *cry*1/*CRY*1, with *cry*1 coupled to *MATa* in DG2 and *CRY*1 coupled to *MATa* in the others. DG3 and DG5 were shown to be homozygous for *HMLa* and *HMRa*; other strains are assumed to have the standard *HMLα HMRa* genotype. Isolation of meiotic products: Haploid segregants were isolated from two clones from each diploid (DG1, 2, 3, 4, 5, 6, 7) by random spore analysis: asci from sporulated cultures were treated with Glusulase (Endo), sonicated and plated onto medium containing cryptopleurine (for DG1) or canavanine (for the other diploids), which selects for haploids because resistance alleles *cry*1 and *can*1 are recessive[8,15]. Haploid segregants were isolated from sporulated cultures of DG8 and DG9 by ascus dissection. Assay of *MATα* segregants: Spore clones with α mating type were identified by replica plating colonies onto a lawn of *a* cells and assaying mating by prototroph formation. Frequency was the number of α colonies per total number of spore clones tested. Numbers in parentheses indicate the number of α colonies observed. Subsequent crosses showed that all α colonies tested were *MATα*. Tests were performed on all α segregants from DG1, DG2, DG8 and DG9; on 3 each from DG4A, DG4B, DG5B, DG6A and DG6B, and on 19 from DG5A. Sporulation: *MATα*2/*MATa* strains DG1, DG2 and DG3 sporulated efficiently (60% for DG3). *matα*2/*MATa* strains DG4 and DG5 sporulated poorly (4% for DG5). *matα*2/*MATa* strains DG6 and DG7 sporulated efficiently due to homozygosity for *sir*1-1 (47% sporulation for DG6). *matα*1/*matα*2 strains DG8 and DG9 sporulated (albeit inefficiently) due to homozygosity for *rme*1, which bypasses the requirement of *MATα*2 and *MATa*1 for sporulation[4].

Regions of homology and non-homology

Our genetic results indicate that *MATα* and *MATa* contain regions of homology as well as regions of non-homology: two sites in *MATα*1 are non-homologous with *MATa*, and two sites in *MATα*2 are homologous with *MATa* (Fig. 1). Studies are in progress to determine whether *MATa*$^+$ recombinants can be obtained from *mata*$^-$/*MATα* diploids and *mata*$^-$ recombinants from MATα$^-$/MATa diploids. Although the extent and position of the homology between *MATα*2 and *MATa* cannot be rigorously determined from our data, the similar frequencies of *MATα*$^+$ recombinants from *matα*2/*MATa* and *matα*2/*matα*1 diploids lead us to favour the view that the homologous sequence is outside the non-homologous sequence (Fig. 1*b*, top) rather than being within it (Fig. 1*b*, bottom). These results therefore suggest that the genetic information determining cell type is larger than the non-homologous sequence observed in heteroduplexes between *MATa* and *MATα* (refs 6, 7). This view is reinforced by the observation that the α mating type locus codes for two transcripts, one primarily within the non-homologous region, and the other primarily outside this region[11,12]. Our observations on homology and non-homology between *MATα* and *MATa* lead us to propose that the latter RNA codes for *MATα*2 and the former for *MATα*1.

Yeast cells can change the information present at *MAT* (an *a* or α 'cassette') by replacing it with a cassette from *HML* or *HMR* (reviewed in ref. 3). This change reflects a transposition of at least the Y region information from *HML* or *HMR* to *MAT* (refs 7, 13). Is the Y region the only information that is transposed? Our findings indicate that homologous sequences are also transposed. This conclusion follows from the observations that MATα mutations can be switched to *MATα*2$^+$ by mating type interconversion[14] and, as shown here, that MATα mutations are outside the non-homology region. Hence, the cassette that is transposed during mating type interconversion must include not only Y information but also flanking sequence information (from X or Z).

We thank Kim Nasmyth and Jim Hicks for communicating their results before publication, Karen Sprague for comments on the manuscript, Julie Dunn for preparation of the manuscript, and David Hagen for preparation of the figure. This work has been supported by Research Career Development Award AI-00163 and USPHS research grant AI-13462 to I. H., by PHS Molecular Biology Training Grant GM-00715 (P. von Hippel, Director), and by American Cancer Society postdoctoral fellowship PF-1282 to G. S.

Received 11 August; accepted 4 November 1980.

1. MacKay, V. L. & Manney, T. R. *Genetics* **76,** 273–288 (1974).
2. Strathern, J. N., Hicks, J. B. & Herskowitz, I. *J. molec. Biol.* (in the press).
3. Herskowitz, I. *et al.* in *The Molecular Genetics of Development* (eds Loomis, W. & Leighton, T.) 79–118 (Academic, New York, 1980).
4. Kassir, Y. & Simchen, G. *Genetics* **82,** 187–202 (1976).
5. Klar, A. J. S., Fogel, S. & Radin, D. N. *Genetics* **92,** 759–776 (1979).
6. Nasmyth, K. A. & Tatchell, K. *Cell* **19,** 753–764 (1980).
7. Strathern, J. N., Spatola, E., McGill, C. & Hicks, J. B. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2839–2843 (1980).
8. MacKay, V. L. & Manney, T. R. *Genetics* **76,** 255–271 (1974).
9. Herskowitz, I., Rine, J., Sprague, G. Jr & Jensen, R. in *Mobilization and Reassembly of Genetic Information* Vol. 17 (eds Scott, W. A., Werner, R., Schultz, J. & Joseph, D. R.) 133–151 (Academic, New York, 1980).
10. Rine, J., Strathern, J. N., Hicks, J. B. & Herskowitz, I. *Genetics* **93,** 877–901 (1979).
11. Klar, A. J. S., Strathern, J. N., Broach, J. R. & Hicks, J. B. *Nature* **289,** 239–244 (1980).
12. Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. & Smith, A. *Nature* **289,** 244–250 (1980).
13. Hicks, J. B. & Herskowitz, I. *Genetics* **85,** 373–393 (1977).
14. Strathern, J. N., Blair, L. C. & Herskowitz, I. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3425–3429 (1979).
15. Skogerson, L., McLaughlin, C. S. & Wakatama, E. *J. Bact.* **116,** 818–822 (1973).

Nucleotide sequences at host–proviral junctions for mouse mammary tumour virus

John E. Majors & Harold E. Varmus

Department of Microbiology and Immunology, University of California, San Francisco, California 94143

Proviruses cloned from rat cells infected with mouse mammary tumour virus, a B-type retrovirus regulated by glucocorticoid hormones, show the structural features of transposable elements: short inverted repeats conclude long direct repeats at the ends of viral DNA, and short sequences of cellular DNA are duplicated during integration and flank each provirus. The integrative mechanism joins a precise site in viral DNA to non-homologous sites in host DNA.

INTEGRATION of a DNA copy of viral RNA into the genome of the host cell is a central event in the life cycle of retroviruses (see Fig. 1; for a review see ref. 1). The viral genome is composed of two identical subunits of single-stranded RNA, each 5–10 kilobases long. On infection, a linear duplex form of viral DNA, slightly longer than a subunit of viral RNA and bearing long terminally repeated units (LTRs)[2–4], is synthesized in the cytoplasm. The LTRs are composed of sequences unique to the 3′ and 5′ ends of viral RNA (U_3 and U_5) and a single copy of a short sequence (R) repeated at the ends of viral RNA, in a manner represented by U_3RU_5 (Fig. 1). Some of the linear unintegrated DNA is carried to the nucleus where a portion at least is converted to closed circular species[5], mostly monomers which contain either one or two copies of the LTR sequence[2,3,6–8]. Integration of viral DNA into the host genome can occur at a wide variety of sites in cellular DNA, possibly at random; the structure of integrated (proviral) DNA is invariably very similar, if not identical, to that of unintegrated linear DNA[9,10], although the immediate precursor of the provirus has not been determined. The primary product of transcription is likely to be equivalent to a subunit of genomic RNA, which can be processed to subgenomic, spliced mRNAs for expression of internal cistrons[1].

Mouse mammary tumour virus (MMTV) is a type B retrovirus with at least three genes (*gag*, *pol* and *env*) on genomic RNA subunits of eight to nine kilobases (see Fig. 1). Although its generally poor infectivity in tissue culture has impeded biochemical study, MMTV seems to replicate like other retroviruses and has the experimental advantages of causing mammary carcinomas in mice and showing dramatic regulation of viral RNA synthesis by glucocorticoid hormones in many types of infected cells (for reviews see refs 11, 12). The mechanisms of carcinogenesis and steroidal regulation by this agent are poorly understood, and an analysis of integration sites of MMTV DNA seemed likely to clarify these aspects and also the general mechanism of integration of retroviral DNA.

We used molecular cloning and DNA sequencing to analyse the structure and integration sites of three proviruses of MMTV DNA introduced into the genome of rat cells (XC cells) by infection in culture. Our analysis demonstrates the remarkable precision of the integrative mechanism and reveals additional striking similarities between the structure of proviruses and that of several transposable elements: short inverted repeats (six base pairs) within the LTRs (1.4 kilobases) terminate each provirus, and duplications of five or six base pairs of host cell DNA are generated during the integration process.

To optimize the possibility of correlating structural and functional properties of proviruses, we isolated a set of clonal rat (XC) cell lines which contain single MMTV proviruses. The lines were chosen from a larger set of clonal lines of XC cells which had been infected by about one infectious unit per cell with the C3H strain of MMTV[13]. To analyse these clones for MMTV sequences, DNA from each clone was digested with *Eco*RI, electrophoresed through agarose, blotted on to nitrocellulose paper by the technique of Southern[14] and hybridized with ^{32}P-MMTV cDNA. Because there is a single recognition site for *Eco*RI within the MMTV(C3H) genome, each normal provirus should yield two hybridizing fragments, the sizes of which are determined by the location of recognition sites in neighbouring cell DNA (Fig. 1*b* shows a restriction map of MMTV DNA).

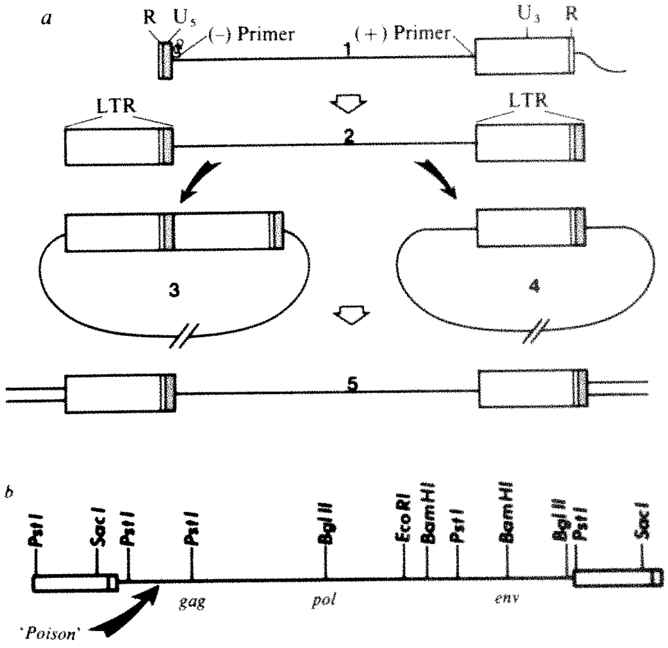


Fig. 1 *a*, Steps in the replication of retroviruses. The RNA subunit of the MMTV genome (1) is shown with features relevant to the present study: the site for priming of the negative (−) DNA strand by tRNA^{Lys} bound near the 5′ terminus[21,22] (shown on the left); the site probably used for priming (+)strand synthesis[2,4]; the short direct repeat (R, probably ~13 bases for MMTV) at both ends of viral RNA; and the regions unique to the 3′ and 5′ ends of viral RNA located between the priming sites and R (U_3, ~1,300 bases; U_5, ~130 bases). At an early step in DNA synthesis these two regions are fused via the R sequence to form the LTR U_3RU_5 which is in turn duplicated during the completion of linear duplex DNA[2–4,6]. The linear DNA (2) is a precursor to the two circular forms which contain either two copies (3) or one copy (4) of the LTR unit[2,3,6,7]. One of these unintegrated forms (2, 3 or 4) is presumably the precursor to the integrated (proviral) form[9,10,23] (5), which serves as the template for synthesis of viral RNA. *b*, A map of restriction sites in MMTV-C3H DNA which are relevant to the present study[6,23]. As described in the text we have been unable to clone the region marked 'poison'.

0028-0836/81/040253—06$01.00

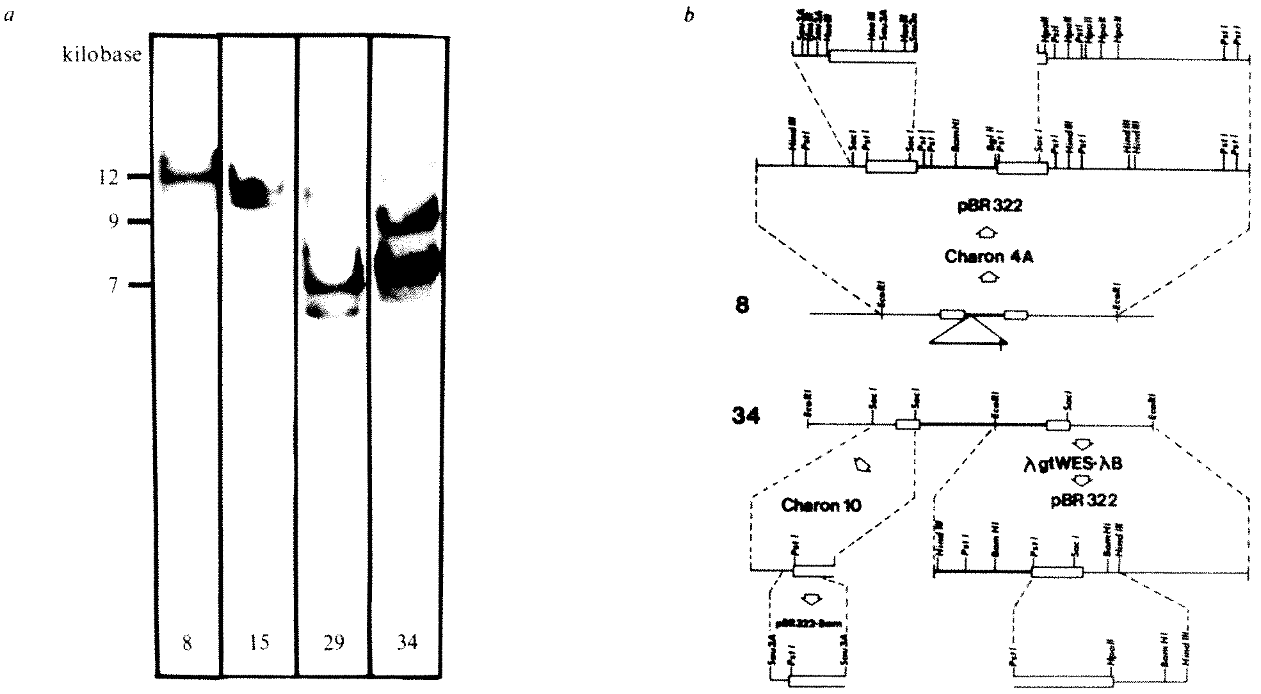


Fig. 2 *a*, Identification of MMTV proviruses in clones of infected rat cells. Rat (XC) cells[46] were infected with MMTV-C3H, seeded in multiple-well plates at a concentration <1 cell per well, and propagated in Dulbecco's minimal essential medium plus 10% calf serum (without added hormones). Purified DNA prepared from each clonal line ($\sim 10\ \mu g$) was digested with *Eco*RI, electrophoresed in 0.8% agarose, transferred to nitrocellulose sheets[14] and hybridized with ^{32}P-labelled cDNA randomly transcribed from MMTV-C3H RNA[6,23]. The figure shows an autoradiogram of an analysis of DNAs from four clones, including 8, 15 and 34, described in the text. *b*, Strategies for cloning the virus–cell junctions of the proviruses from lines 8 and 34. DNA (mg aliquots) from each line were cleaved with *Eco*RI and passed through 0.8% agarose in a fractionating gel electrophoresis apparatus[47]. The eluted fractions were analysed for fragments of interest as described for *a*. To isolate the junctions from line 8, DNA from the fractions containing the 12-kilobase *Eco*RI fragment was ligated with purified Charon 4A arms and then packaged *in vitro* as described by Blattner *et al.*[48]. The assembled phages were plated onto the strain DP50SupF (ref. 18) and screened for proviral inserts by the procedure of Benton and Davis[49] using MMTV ^{32}P-labelled cDNA$_{\mathrm{rep}}$. A positively hybridizing plaque was isolated for further analysis. The insert from this phage was transferred into pBR322 using vector DNA that had been cleaved with *Eco*RI and treated with calf intestinal alkaline phosphatase. DNA from one such clone was isolated and used as a substrate for restriction cleavage analysis. For sequencing, the indicated *Sac*I fragments were purified by preparative electrophoresis through 0.8% Sea-Plaque (Marine Colloids) agarose gels, then further restricted for the construction of fine structure maps. The two junctions from the provirus in line 34 were isolated separately. DNA from fractions containing the right end 4.4-kilobase *Eco*RI fragment was ligated with purified arms from the vector λgtWES · λB, then packaged, plated and assayed as described above. After isolation, the purified insert was then transferred into pB322 and used in restriction enzyme analysis. DNA from fractions containing the six-kilobase *Eco*RI fragment from the left end of the provirus was further digested with *Sac*I, then ligated to purified arms from the vector Charon 10 (ref. 20). Packaging, plating and screening were as described above. After preliminary characterization and preparative electrophoresis through 0.8% Sea-Plaque agarose, the inserted *Sac*I fragment was cleaved with the restriction enzyme *Sau*3A, and the cleavage products ligated to *Bam*HI-cleaved pBR322 that had been treated with calf intestinal alkaline phosphatase. After transformation of the bacterial strain HB101 (ref. 50), ampicillin-resistant recombinants were screened with a probe specific for the MMTV LTR, using the procedure of Hanahan and Meselson[51], for those containing the junction.

Using this criterion, of the 45 clones analysed, 12 had two or more proviruses, 14 had a single provirus and 19 had escaped infection. Figure 2*a* shows the results of such an analysis of four of the clones carrying single proviruses. We chose three of these, designated 8, 15 and 34, for further analysis. One (clone 8) carried a truncated provirus (see Fig. 2*b*) expressed under the control of glucocorticoid hormones (our unpublished data and those of D. Robertson); the second (clone 34) contained a full-sized provirus from which little, if any, stable RNA was transcribed in the presence or absence of hormone (Fig. 2*b*, unpublished data); the third, not fully analysed here, contained a full-sized provirus under steroid control (unpublished data).

Cloning unintegrated and proviral MMTV DNA

Our choice of strategies for cloning MMTV proviral DNA was influenced by a difficulty we encountered in attempting to isolate unintegrated forms of MMTV DNAs by molecular cloning in *Escherichia coli* (similar difficulties have been encountered in attempts to clone integrated MMTV DNA (ref. 15 and D. Ucker, personal communication)). Circular MMTV molecules were isolated from dexamethasone-stimulated rat XC cells that were chronically infected with the C3H strain of MMTV. (Heterologous cells chronically infected by MMTV generally contain 1–10 copies of unintegrated DNA, apparently synthesized by reverse transcriptase, when grown in the presence of glucocorticoids[6,16,17].) After partial purification by Hirt fractionation and CsCl–propidium diiodide equilibrium centrifugation[6,16], the viral DNA was cleaved with *Eco*RI and cloned into the vector λgtWES·λB (ref. 18). Although the virus-specific material that we ligated to the vector arms seemed to be primarily intact *Eco*RI-cleaved circular molecules, each of the positively hybridizing λgtWES · MMTV recombinants had suffered deletions or rearrangements of MMTV sequences. The deletion end points were variable, but each lesion covered a common set of MMTV sequences that lie within a region of the genome (which is designated 'poison' in Fig. 1) which probably codes for gag proteins.

To investigate this phenomenon further, we cleaved the same DNA preparation with *Pst*I, an enzyme with several recognition sites within the MMTV genome, and cloned the fragments in the *Pst*I site of the plasmid vector pBR322. We isolated recombinants containing all the expected MMTV *Pst*I fragments except the 1-kilobase fragment containing the 'poison' sequence. We concluded from these experiments that there is a sequence within the 1-kilobase *Pst* I fragment which propagates poorly, if at all, in the strains of *E. coli* used. We cannot explain this result and have not yet isolated those sequences. In analysing integration sites, we were forced to use strategies which allowed us to circumvent this cloning difficulty.

Figure 2*b* shows partial restriction maps of the two proviruses and describes the cloning strategies. The provirus from clone 34 seems normal, that is, the restriction fragments generated from internal regions of the provirus were identical to those of unintegrated linear DNA. The provirus in clone 8 is abnormal in that cleavage with *Eco*RI released only a single 12-kilobase fragment which hybridized with MMTV cDNA$_{rep}$. Further analysis with *Pst*I, *Bam*HI, *Hin*dIII and *Sac*I showed that the clone 8 provirus had incurred a large deletion of viral sequences, probably covering the *gag–pol* region of the MMTV genome; this deletion was fortuitous in that it eliminated both the poison sequence and the internal *Eco*RI site, and we were able to clone the entire steroid-responsive provirus with its flanking cell sequences in the *Eco*RI site of the λ phage vector Charon 4A (ref. 19).

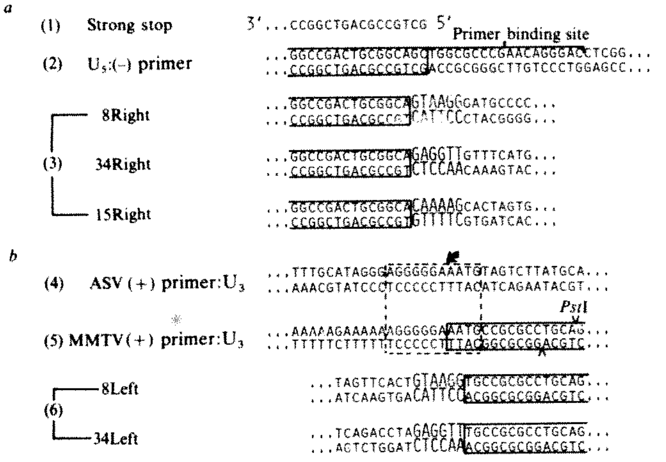

a
(1) Strong stop
3'...CCGGCTGACGCCGTCG 5'
Primer binding site
(2) U$_5$:(–) primer
...GGCCGACTGCGGCAGCTGGCGCCCGAACAGGGACCTCGG...
...CCGGCTGACGCCGTCGACCGCGGGCTTGTCCCTGGAGCC...
(3)
8Right
...GGCCGACTGCGGCAGTAAGGGATGCCCC...
...CCGGCTGACGCCGTCATTCCCTACGGGG...
34Right
...GGCCGACTGCGGCAGAGGTTGTTTCATG...
...CCGGCTGACGCCGTCTCCAACAAAGTAC...
15Right
...GGCCGACTGCGGCACAAAAGCACTAGTG...
...CCGGCTGACGCCGTGTTTTCGTGATCAC...
b
(4) ASV (+) primer:U$_3$
...TTTGCATAGGGAGGGGGAAATGTAGTCTTATGCA...
...AAACGTATCCCTCCCCCTTTACATCAGAATACGT...
PstI
(5) MMTV(+) primer:U$_3$
...AAAAAGAAAAAAGGGGGAAATGCCGCGCCTGCAG...
...TTTTTCTTTTTTCCCCCTTTACGGCGCGGACGTC...
(6)
8Left
...TAGTTCACTGTAAGGTGCCGCGCCTGCAG...
...ATCAAGTGACATTCCACGGCGCGGACGTC...
34Left
...TCAGACCTAGAGGTTTGCCGCGCCTGCAG...
...AGTCTGGATCTCCAAACGGCGCGGACGTC...

Fig. 3 *a*, Junctions at the right ends. The end of U$_5$ was determined from the sequence of 'strong stop' (–) strand DNA (RU$_5$) synthesized *in vitro* (1) and from the sequence of the junction between the left-hand LTR and the region encoding the binding site for tRNA^{Lys} (2). The latter was derived from a *Pst*I subclone of the relevant region from line 8 provirus. These sequences were then compared with the sequences of the junctions between viral and host DNA at the right ends of proviruses from lines 8, 15 and 34 (3). The sequence of strong stop DNA was determined by two methods: a modification of the di-deoxy chain termination method[52] was used to analyse products synthesized by purified avian myeloblastosis virus (AMV) DNA polymerase on a template of MMTV-C3H 70S RNA; the chemical sequencing technique of Maxam and Gilbert[53] was used to analyse strong stop DNA synthesized by detergent-disrupted virions of MMTV-GR (provided by Dr Nurul Sarkar). For each junction sequence, both viral: cell and U$_5$:primer binding site, DNA was labelled at a *Hpa*II site within the U$_5$ sequence, 30 base pairs from its right end. *b*, Junctions at the left ends. The end of the U$_3$ sequence was estimated from the position of the indicated *Pst*I site and from a comparison of sequences from the Schmidt-Ruppin A strain of ASV DNA[24,25] (4) and from MMTV line 8 proviral DNA at the leftward boundary of the right-hand LTR (5) (see text). The 11-base homology is boxed in and the arrow indicates the boundary of the ASV U$_3$ region. These sequences were then compared with those from the viral-host junctions at the left ends of the two proviruses (6). The (+) primer:U$_3$ junction was derived from DNA labelled at the *Bgl*II site within the line 8 proviral sequences. The cell-virus junctions were derived from DNA labelled at sites within the host flanking sequences, a *Hae*III site for line 8 and a *Sau*3A site for line 34 (see Fig. 2*b*). Presumed viral U$_5$ and U$_3$ sequences are enclosed within solid lines and the repeated cell sequences are indicated by large letters. Sequences were obtained by the chemical cleavage method of Maxam and Gilbert[53], using DNA fragments that had been labelled at 3' ends either with AMV reverse transcriptase or *E. coli* DNA polymerase I.

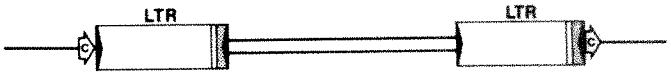


Fig. 4 Schematic diagram of an MMTV provirus. The inverted repeats at the ends of the LTRs are shown as inverted solid triangles. The direct repeat of the cell sequence at the junction site is indicated by C within the open arrow. Viral DNA is shown as a double line, flanking cellular DNA as a single line. Other symbols are as described for Fig. 1.

The strategy for cloning the ends of the proviruses in clones 34 and 15 was dictated by the need to circumvent the poison sequence. The *Eco*RI fragments containing the right ends were cloned in λgtWES · λB. The *Sac*I site in the U$_3$ region of the LTR at the left end separates the virus–cell junction from the poison sequence; therefore the left-hand virus–cell junction from line 34 was cloned as a *Sac*I fragment in the lambda phage vector Charon 10 (ref. 20) (Fig. 2*b*).

Determining the boundaries of the LTR

To determine the fate of viral sequences during integration we had to know accurately the structure of viral DNA before integration. It was particularly necessary to define the boundaries of the long terminal repeat, the right and left ends of which are also the right and left ends of the linear form of unintegrated DNA (Fig. 1).

The right end should be the site of initiation of viral reverse transcription[21], a sequence determined by direct DNA sequencing of the initial or 'strong stop' (–) strand DNA made *in vitro* (unpublished). Figure 3*a* shows the sequence of the region in the provirus from clone 8 covering the boundary between the left LTR and viral sequences outside the LTR, that is, the boundary between U$_5$ (the strong stop sequence) and the tRNA primer binding site. The strong stop transcript begins GCTGC, which is immediately preceded by 18 bases identical to the 3' end of tRNA$_3^{Lys}$, the primer for MMTV negative strand synthesis[22].

Present replication models argue that the left end of the LTR should correspond to the initiation point for (+) strand strong stop synthesis[4]. We have no direct evidence with which to define this point; therefore we have resorted to indirect arguments. There is a *Pst*I site within the LTR, near the left end of this previously described repeat unit of ~1.4 kilobases (see Fig. 1 and refs 6, 10, 23). Figure 3*b* shows the sequence extending leftwards from that *Pst*I site, towards the viral *Bgl*II site, which is on the 5' side of the rightward LTR. This sequence should contain the left boundary of the LTR, and within it are two features which seem useful for defining the boundary: (1) in the (+) strand there is a polypurine tract of 20 bases ending ~10 bases from the *Pst*I site; (2) a sequence of 11 base pairs, including the last nine of the polypurine tract, has precise homology with a sequence from the analogous region of avian sarcoma virus (ASV) DNA [24,25]. Swanstrom *et al.*[24] have shown, by sequencing the junction between the LTRs of a cloned ASV circular molecule containing two copies of the LTR, that the left end of the ASV LTR is probably at the position marked with an arrow in Fig. 3*b*. We suggest that the left end of the MMTV LTR is at the homologous site and that the polypurine tract is involved with plus strand strong stop priming. Comparison of the proposed sequences at the left and right ends of the LTR unit show them to be related by an inverted repeat of six bases (AATGCCGC.... GCGGCAGC), starting two base pairs from the ends. This supports our choice of boundaries because short inverted repeats of various lengths have been found at the ends of the LTRs of murine leukaemia and sarcoma viruses[26-28], spleen necrosis virus (SNV)[29] and ASV (ref. 24).

Sequences at the host–viral junctions

Sequencing of virus–cell junctions was facilitated by transferring the *Eco*RI fragments from Charon 4A·8, λgtWES·34R and λgtWES·15R into the plasmid pBR322. To analyse the cell–virus junction in Charon 10·34L, the *Sac*I insert was cleaved with *Sau*3A and the fragment containing the junction cloned into the *Bam*HI site of pBR322. Restriction maps derived from certain of these plasmid clones are shown in Fig. 2. Sequences from the right-end junctions were obtained by labelling at a *Hpa*II site in the U_5 region and sequencing across the junction into host DNA. Figure 3 compares those sequences with that from the boundary between the right end of the leftward LTR and the tRNA primer binding site. Two proviruses are missing at least one base of U_5 sequence and the right end of the third provirus (from line 15) is missing at least two bases from the U_5 sequence. (All three probably lack the first two bases of U_5 at this end.) Also shown are the viral–host junction sequence derived from the 34L *Sau*3A subclone and that of the left junction of the provirus from line 8 derived from the *Eco*RI fragment cloned into pBR322. A comparison with the expected left end of the LTR shows that each provirus is lacking two base pairs from the proposed terminus of the LTR.

Comparison of the cellular sequences at the right and left junctions of the provirus from both lines 8 and 34 shows that there are direct repeats in the cellular sequence at the site of integration—these comprise six base pairs for both proviruses. In each case, the assignment of a GC base pair to cellular or viral sequences is ambiguous. However we argue that symmetry and consistency make a six-base pair repeat of cell DNA, with two base pairs missing from both ends of viral DNA, the logical provisional assignment.

The two base pairs removed are those which are mismatched in the inverted repeats at the ends of the LTR, and their removal leaves a perfect inverted repeat of 6 base pairs in the viral sequences at the ends of the provirus. Thus an MMTV provirus has the structure shown in Fig. 4: unique coding sequences, flanked by large terminal direct repeats, which in turn are flanked by short inverted repeats, with a direct repeat in the cell sequence at the integration site. These features are strikingly analogous to those of several transposable elements, particularly Tn9 of *E. coli*[30,31], *copia* of *Drosophila*[32-34] and Ty1 of yeast[35-37].

The repeated cell sequence flanking the provirus is duplicated due to integration

It has been shown for several transposable elements that the direct repeat in the host sequence which flanks the element is the result of a duplication generated during integrative recombination[38]. To verify that a similar process is involved in retrovirus integration and to exclude the formal possibility of integration into a duplicated region of host DNA, we cloned the *Eco*RI fragment into which the provirus from line 8 was integrated. A restriction map of that fragment is indistinguishable from that of the flanking sequences in the fragment containing the line 8 provirus (Fig. 5). The sequence of the integration site shows that the six-base pair repeat which flanks the provirus in clone 8 was present only once in the integration site before insertion; thus we conclude that the duplication was generated during the integration process.

Specificity of integration sites

Analyses of proviruses of several retroviruses indicate that there must be many integration sites available for each virus in each host cell[9,10,23,39] (see Fig. 2). However, mapping procedures are inadequate to determine whether a repeated host sequence is common to the integration sites used by any of these viruses. Figure 6 shows a comparison of the sequences for the integration sites for MMTV proviruses in clones 8 and 34. There seem to be no common features between these sites or between either site and the ends of the proviruses. In addition, there is no obvious relationship between these sequences and those flanking the right end of the provirus in clone 15 (Fig. 3*a*) or those flanking the right ends of three proviruses present in a mouse mammary tumour cell[15]. This limited survey is thus consistent with the idea that integration occurs randomly within the host genome; however, our small sample size makes the detection of any subtle specificities unlikely. For instance, fine structure analysis of the target sites for a large number of insertions of the transposable elements Tn3 and Tn9 shows a correlation between the probability of integration and both base composition and homology with end sequences[38].

Models for integration of retroviral DNA

This analysis of MMTV DNA and similar analyses of other retroviral DNAs[27-29] argue that any model for retroviral integration should account for the following observations: (1) A precisely defined region of viral DNA, corresponding within a few base pairs to the ends of linear DNA, seems to be invariably linked to host DNA. This implies either that a mechanism exists to recognize a specific site in viral circular DNA or that linear DNA integrates directly. (2) A limited number of base pairs, probably two, are lost from both the right and left ends of viral DNA during integration. Note that these bases are 'lost' only with respect to the probable structure of the unintegrated linear DNA. The proximal precursor to the provirus could be a circular form with a different complement of sequences in the relevant region. Furthermore, the lost base pairs do not affect the genetic

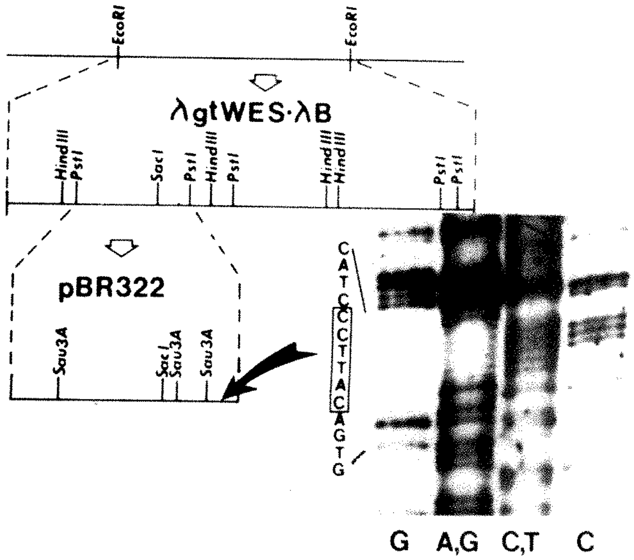


Fig. 5 The target site for the provirus in line 8. A hybridization probe made against the entire 12-kilobase *Eco*RI fragment from the line 8 provirus cloned in pBR322 was used to isolate fractions containing the target site fragment from a preparation of fractionated *Eco*RI-cleaved DNA from line 34. DNA from those fractions was ligated to purified λgtWes · λB arms and a recombinant phage carrying the fragment which contained the target site was isolated as described in Fig. 2 legend. The purified recombinant DNA was used for restriction enzyme cleavage analysis. For sequence analysis the *Eco*RI insert was first transferred into pBR322; the *Pst*I fragment containing the target site was then sub-cloned into the *Pst*I site of pBR322. The sequence was obtained by Maxam–Gilbert chemical cleavage analysis of the *Sau*3A-*Pst*I fragment which had been labelled at the 3' terminus of the *Sau*3A site. The relevant part of a 10% acrylamide sequencing gel and the inferred sequence are shown. The band present in all four lanes is the result of a partial *Sau*3A cleavage 5' to the sequence CATC.

```
8     ...GTAGCAAGGGGCCTAGTTCACTGTAAGGGATGCCCCAGACCA...
      ...CATCGTTCCCCGGATCAAGTGACATTCCCTACGGGGTCTGGT...

34    ...AAGTGCCCCACTTTCAGACCTAGAGGTTGTTTCATGATTXCA...
      ...TTCACGGGCTGAAAGTCTGGATCTCCAACAAAGTACTAAXGT...
```

Fig. 6 A comparison of two targets for MMTV integration into rat XC cell DNA. The sequence of the integration site used for the provirus in line 8 was determined directly as described in Fig. 5 legend. The sequence of the site used in line 34 was deduced from the sequences flanking the ends of the provirus (Fig. 3).

capacity of the provirus because they originate from regions which are repeated in linear DNA but present uniquely in viral RNA (see Fig. 1). (3) Host sequences are duplicated at the site of integration. (4) Viral DNA terminates with short inverted repeats. (5) A large number of regions in host DNA, including sites with sequences unrelated to each other or to viral DNA, can accommodate proviral DNA.

Although proviral elements from all retroviruses analysed have shown these general features, there are potentially significant differences in their fine structures: (1) The terminal base pairs which are lost from MMTV DNA after integration are not included within terminal inverted repeat sequences, whereas those from other retroviral DNAs are part of an inverted repeat[27–29]. (2) The size of the inverted repeats in proviral DNA varies from 20 base pairs (with four mismatches) in Moloney murine leukaemia virus (Mo-MuLV) DNA[28] to three base pairs in SNV DNA[29]. (3) The length of duplication in cellular sequences generated by integration seems to be specific for each retrovirus; thus, Moloney strains of MuLV and murine sarcoma virus (MSV) generate a four-base pair repeat[27,28], SNV generates a five-base pair repeat[29] and MMTV probably generates a six-base pair repeat.

The duplication of cellular sequences strongly suggests that an early step in the integrative process will prove to be the introduction of a staggered scission in host chromosomal DNA, as proposed for transposable elements[40,41]. The lack of homology between the two MMTV integration sites described here indicates that the enzyme responsible for this nucleolytic activity may have little sequence specificity. No such host or virus-coded activity has yet been described, although an endonucleolytic activity, encoded in the viral *pol* gene, has been found in virions of avian myeloblastosis virus[42].

Unfortunately, our structural studies shed little light on the identity of the viral integrative precursor. Both the linear and the two circular forms of unintegrated DNA can be accommodated as precursors within credible models for integration[28], each of which is closely related to that proposed by Shapiro[41] for the transposition of bacterial transposons. Recent work on the mechanism of restriction of MuLV replication by alleles at the *Fv-1* locus can be interpreted as supporting a circular precursor[43,44]. In some cases, *Fv-1* alleles inhibit both the accumulation of circular forms of MuLV DNA and the integration of MuLV DNA into mouse DNA. However, it is possible that these two observations both reflect the defective nature of linear DNA; thus the identity of the proximal precursor to the provirus remains an open issue.

Functional implications of proviral structure

The significance of the structural similarities between proviruses and transposable elements (Fig. 4) is uncertain in the absence of mechanistic details about retrovirus integration or about transposition of bacterial or eukaryotic transposons. Duplications at the sites of integration argue that the generation of staggered single-stranded scissions in host DNA is an early event in both processes. Terminal inverted repeats should allow an enzyme with a single recognition specificity to operate on both ends of an integrating molecule; however, this feature may not be essential for transposition because the Ty1 elements of yeast lack terminal inverted repeats[36,37].

There is no information concerning the possibility that proviruses can be relocated at new sites in host genomes by means other than transcription into RNA, subsequent reverse transcription and reintegration. Experiments to test for direct transposition are difficult to design and unlikely to yield unambiguous results. However, there is evidence to suggest that retroviruses can show other forms of behaviour observed for transposable elements, including deletion formation, insertional mutagenesis and promotion of transcription into flanking cellular DNA[45]. Each of these may have significant genetic consequences for the host, apart from the expression of viral genes.

The structural studies reported here do not directly further our understanding of hormonal regulation or carcinogenesis in MMTV-infected cells. However, the provirus from clone 8 is under glucocorticoid control and has been cloned intact with large regions of flanking cellular DNA; it should serve as a useful reagent for examining the role of viral and context (host) sequences in the mediation of the hormone response. The sequence of the MMTV LTR (studies to be presented elsewhere) includes an A+T-rich region characteristic of putative eukaryotic promoters. This region, which reads TATAAAAGA, is positioned 24–32 bases from the probable initiation site for RNA synthesis, the site corresponding to the 5′ terminus of genomic RNA. Thus we favour the notion that LTRs, like repeated units in transposons, may provide signals for gene expression[24–29].

We thank Bertha Cook for secretarial assistance. This work was supported by USPHS grants CA 12705, CA 19289, Training Grant 1T32 Ca 09043 and American Cancer Society VC-70. J.E.M. was supported by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research.

Received 28 August; accepted 13 November 1980.

1. Bishop, J. M. *A. Rev. Biochem.* **47,** 35–88 (1978).
2. Shank, P. R. *et al. Cell* **15,** 1383–1396 (1978).
3. Hsu, T. W., Sabran, J. L., Mark, G. E., Guntaka, R. V. & Taylor, J. M. *J. Virol.* **28,** 810–818 (1978).
4. Gilboa, E., Mitra, S. W., Goff, S. & Baltimore, D. *Cell* **18,** 93–100 (1979).
5. Shank, P. R. & Varmus, H. E. *J. Virol.* **25,** 104–114 (1978).
6. Shank, P. R., Cohen, J. C., Varmus, H. E., Yamamoto, K. R. & Ringold, G. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2112–2116 (1978).
7. Yoshimura, F. K. & Weinberg, R. A. *Cell* **16,** 323–332 (1979).
8. Kung, H. J., Shank, P. R., Bishop, J. M. & Varmus, H. E. *Virology* **103,** 425–433 (1980).
9. Hughes, S. H. *et al. Cell* **15,** 1397–1410 (1978).
10. Ringold, G. M., Shank, P. R., Varmus, H. E., Ring, J. & Yamamoto, K. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 665–669 (1979).
11. Moore, D. H. *et al. Adv. Cancer Res.* **29,** 347–418 (1979).
12. Varmus, H. E., Ringold, G. & Yamamoto, K. R. in *Glucocorticoid Hormone Action* (eds Baxter, J. D. & Rousseau, G. G.) 253–278 (Springer, Berlin, 1979).
13. Fine, D. L., Plowman, J. K., Kelley, S. P., Arthur, L. O. & Hillman, E. A. *J. natn. Cancer Inst.* **52,** 1881–1886 (1974).
14. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
15. Hager, G. L. & Donehower, L. A. in *Animal Virus Genetics: ICN–UCLA Symp. molec. cell. Biol.* (eds Fields, B., Jaenisch, R. & Fox, C. F.) (Academic, New York, in the press).
16. Ringold, G. M., Yamamoto, K. R., Shank, P. R. & Varmus, H. E. *Cell* **10,** 19–26 (1976).
17. Ringold, G. M., Shank, P. R. & Yamamoto, K. R. *J. Virol.* **26,** 93–101 (1978).
18. Leder, P., Tiemeier, D. & Enquist, L. *Science* **196,** 175–177 (1977).
19. Blattner, F. R. *et al. Science* **196,** 161–169 (1977).
20. deWet, J. R. *et al. J. Virol.* **33,** 401–410 (1980).
21. Taylor, J. M. *Biochim. biophys. Acta* **473,** 57–72 (1977).
22. Peters, G. & Glover, C. C. *J. Virol.* **35,** 31–40 (1980).
23. Cohen, J. C., Shank, P. R., Morris, V. L., Cardiff, R. & Varmus, H. E. *Cell* **16,** 333–346 (1979).
24. Swanstrom, R., DeLorbe, W., Bishop, J. M. & Varmus, H. E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
25. Czernilofsky, A. P., DeLorbe, W., Swanstrom, R., Varmus, H. E. & Bishop, J. M. *Nucleic Acids Res.* **8,** 2967–2984 (1980).
26. Sutcliffe, J. G., Shinnick, T. M., Verma, I. M. & Lerner, R. A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3302–3306 (1980).
27. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3937–3941 (1980).
28. Shoemaker, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 3932–3936 (1980).
29. Shimotohno, K., Mizutani, S. & Temin, H. M. *Nature* **285,** 550–554 (1980).
30. MacHattie, L. A. & Jackowski, J. B. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 219–228 (Cold Spring Harbor Laboratory, New York, 1977).
31. Johnsrud, L., Calos, M. & Miller, J. H. *Cell* **15,** 1209–1219 (1978).
32. Strobel, E., Dunsmuir, P. & Rubin, G. M. *Cell* **17,** 429–439 (1979).
33. Dunsmuir, P., Brorein, W. J. Jr, Simon, M. A. & Rubin, G. M. *Cell* **21,** 575–579 (1980).
34. Levis, R., Dunsmuir, P. & Rubin, G. M. *Cell* **21,** 581–588 (1980).
35. Cameron, J. R., Loh, E. Y. & Davis, R. W. *Cell* **16,** 739–751 (1979).
36. Farabaugh, P. J. & Fink, G. R. *Nature* **286,** 352–356 (1980).
37. Gafner, J. & Phillipsen, P. *Nature* **286,** 414–418 (1980).
38. Calos, M. P. & Miller, J. H. *Cell* **20,** 579–595 (1980).
39. Steffen, D. & Weinberg, R. A. *Cell* **15,** 1003–1010 (1978).

40. Grindley, N. D. F. & Sherratt, D. J. *Cold Spring Harb. Symp. quant. Biol.* **43,** 1257–1268 (1978).
41. Shapiro, J. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1933–1937 (1979).
42. Golomb, M. & Grandgenett, D. P. *J. biol. Chem.* **254,** 1606–1613 (1979).
43. Yang, W. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 2994–2998 (1980).
44. Jolicouer, P. & Rassart, E. *J. Virol.* **33,** 183–195 (1980).
45. Majors, J. E. *et al. Cold Spring Harb. Symp. quant. Biol.* **45** (in the press).
46. Svoboda, J. *Nature* **186,** 980–981 (1960).
47. Leder, P. *et al. Cold Spring Harb. Symp. quant. Biol.* **42,** 915–920 (1978).
48. Blattner, F. R. *et al. Science* **202,** 1279–1288 (1978).
49. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
50. Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41,** 459–472 (1969).
51. Hanahan, D. & Meselson, M. *Gene* **10,** 63–68 (1980).
52. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463–5467 (1977).
53. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).

Nucleotide sequence and formation of the transforming gene of a mouse sarcoma virus

Charles Van Beveren, Janice A. Galleshaw, Vivian Jonas, Anton J. M. Berns*, Russell F. Doolittle†, Daniel J. Donoghue & Inder M. Verma

Tumor Virology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138
†Department of Chemistry, University of California, San Diego, La Jolla, California 92093

The complete nucleotide sequence of the transforming gene of a mouse sarcoma virus has been determined. It codes for a protein of 374 amino acids. The nucleotide sequence of the junctions between a murine leukaemia virus and cellular sequences leading to the formation of the viral transforming gene have also been elucidated. The viral transforming sequence and its cellular homologue share an uninterrupted stretch of 1,159 nucleotides, with few base substitutions. The predicted amino acid sequence of the mouse sarcoma virus transforming gene was found to share considerable homology with the proposed amino acid sequence of the avian sarcoma virus oncogene (src) *product.*

MANY of the retroviruses capable of cellular transformation have acquired a normal cellular sequence as an integral part of their genome[1]. The acquired cellular sequences can be expressed either as an independent viral gene or as a product fused with viral structural proteins[2-10]. With the exception of Rous sarcoma viruses (RSV), acquisition of cellular sequences occurs at the expense of viral genomic sequences required for replication[1,11-18]. Not surprisingly, most retroviruses capable of inducing cellular transformation are unable to propagate without helper viruses.

We were interested in the biogenesis of a transforming gene at the level of nucleotide sequences. We chose to study Moloney mouse sarcoma virus (Mo-MSV) which is able to transform fibroblasts *in vitro* and induce neoplasia *in vivo*, but is unable to replicate[19-21]. It was obtained from sarcomas induced in BALB/c mice after passage of high doses of Moloney murine leukaemia virus (Mo-MLV)[22]. The genome of a cloned isolate of Mo-MSV (clone 124, ref. 23) has been characterized extensively. Briefly, the Mo-MSV genomic RNA is about 5.3 kilobases long and shares 60–70% sequence homology with Mo-MLV[24]. Figure 1 shows a schematic heteroduplex formed between a genome-length Mo-MLV cDNA transcript and Mo-MSV 124 genomic RNA. The Mo-MSV genome has a sequence substitution of about 1.2 kilobases and two deletions of approximately 1.2 and 1.6 kilobases. About 2.0 kilobases at the 5′ end and 0.6–0.7 kilobases at the 3′ end are shared between the two viruses[11-15].

We have molecularly cloned the unintegrated circular forms of Mo-MSV clone 124 and Mo-MLV DNA in bacteriophage λ (refs 25, 26). by using molecularly cloned Mo-MSV^{src}-specific DNA fragments as probe, a 14.0-kilobase *Eco*RI fragment from uninfected mouse cells was identified and subsequently molecularly cloned in bacteriophage λ (refs 27, 28). The cloned cellular DNA fragment contains an uninterrupted stretch of about 1.1 kilobases homologous to Mo-MSV-specific sequences. The Mo-MSV^{src}-specific sequences will be referred to here as v-*mos*$^{\mathrm{Mo}}$ [viral (v), mouse sarcoma virus (*mos*), Moloney (Mo)] and their cellular homologue is referred to as c-*mos*$^{\mathrm{Mo}}$ [cellular (c)]. We were interested in determining the nucleotide sequence in the region of the junctions of c-*mos*$^{\mathrm{Mo}}$ and Mo-MLV sequences to learn more about the process leading to the formation of the v-*mos*$^{\mathrm{Mo}}$ gene. We present here the following results: (1) the complete nucleotide sequence of the v-*mos*$^{\mathrm{Mo}}$ gene of Mo-MSV clone 124; (2) the predicted amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product; (3) nucleotide sequences of the 5′ and 3′ extremities of c-*mos*$^{\mathrm{Mo}}$ and Mo-MLV leading to the formation of the v-*mos*$^{\mathrm{Mo}}$ gene; (4) nucleotide sequence comparison of the portions of v-*mos*$^{\mathrm{Mo}}$ and c-*mos*$^{\mathrm{Mo}}$; (5) comparison of the predicted amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product with the proposed avian sarcoma virus oncogene (*src*) product.

Construction of plasmids

The following plasmids were constructed for the purpose of determining the nucleotide sequence of the appropriate regions (Fig. 2). The plasmid, p·MSV-1, contains the entire Mo-MSV DNA, whereas the plasmid, p·MSV-21, contains a majority of the v-*mos*$^{\mathrm{Mo}}$ gene sequences. The cellular DNA fragment containing the c-*mos*$^{\mathrm{Mo}}$ sequence was cloned in bacteriophage Charon 4A and portions of it were subsequently subcloned in the plasmid pBR322. The plasmid, p·CS-1, contains the 5′ half of the c-*mos*$^{\mathrm{Mo}}$ and cellular sequences, whereas p·CS-2 contains the entire coding region of c-*mos*$^{\mathrm{Mo}}$ and adjacent cellular sequences. The plasmid, p·MLV-1, contains the entire Mo-MLV genome. The approximate position of the junction of Mo-MLV and v-*mos*$^{\mathrm{Mo}}$ sequences was obtained from the heteroduplex map (Fig. 1) and physical maps generated by restriction endonucleases[29,30].

Nucleotide sequences of the v-*mos*$^{\mathrm{Mo}}$ gene

The Maxam–Gilbert method of DNA sequencing was used. In general, 5′→3′ sequences were determined by labelling with polynucleotide kinase and [γ-^{32}P]ATP, and the 3′→5′ sequences were determined by labelling the 3′ end with [α-^{32}P]CTP and terminal transferase. The complete nucleotide

---

* Present address: Laboratory of Biochemistry, University of Nijmegen, Nijmegen, The Netherlands.

0028-0836/81/040258—05$01.00

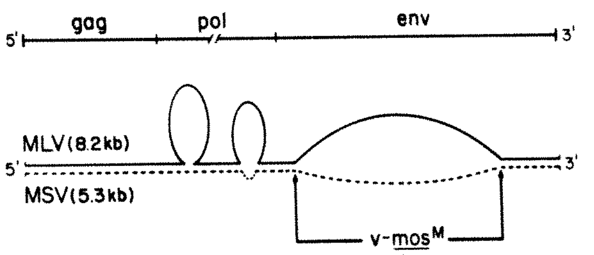


Fig. 1 Schematic heteroduplex formed between genome-length Mo-MLV DNA and Mo-MSV 124 genomic RNA[11]. kb, Kilobases.

sequence of a DNA fragment encompassing the v-*mos*$^{\mathrm{Mo}}$ gene of Mo-MSV 124 is shown in Fig. 3. The nucleotide sequence from just before the first *Bgl*I site through the *Hin*dIII site (1,356 base pairs) was obtained from p·MSV-21 (Fig. 2). The nucleotide sequence of v-*mos*$^{\mathrm{Mo}}$ past the *Hin*dIII site was determined from p·MSV-1.

There is only one major region in which a reading frame in either direction is open for translation. It starts with an ATG about 10 nucleotides downstream from the single *Xba*I site (Fig. 2), and can encode a protein containing 374 amino acids. The open reading frame terminates with an opal codon (TGA) within the acquired cellular sequences. Figure 3 also shows the predicted amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product. Based on the predicted sequence, the largest possible v-*mos*$^{\mathrm{Mo}}$ gene product has a molecular weight of 40,992, which is in good agreement with the estimated size of a 37,000 molecular weight (37 K) protein and a related family of proteins synthesized in an *in vitro* translation of Mo-MSV clone 124 genomic RNA[31]. The synthesis of the 37K protein and a related family of polypeptides is specifically inhibited when the Mo-MSV 124 genomic RNA is preannealed to v-*mos*$^{\mathrm{Mo}}$-specific DNA fragments (J. Papkoff, T. Hunter and I. M. V., unpublished results). The synthesis of other viral proteins, like the p60 *gag* gene product, is not inhibited by the v-*mos*$^{\mathrm{Mo}}$-specific fragments. Note, however, that the 37K and related polypeptides synthesized *in vitro* have not been conclusively shown to be *in vivo* v-*mos*$^{\mathrm{Mo}}$ gene products, as no antisera against these polypeptides are yet available. Based on its amino acid sequence, the v-*mos*$^{\mathrm{Mo}}$ gene product has a net positive charge, in agreement with the mobility of the *in vitro* 37K protein on isoelectric focusing gels (J. Papkoff and T. Hunter, personal communication).

Nucleotide sequence of the c-*mos*$^{\mathrm{Mo}}$ and Mo-MLV junctions

To determine the nucleotide sequences at the 5' and 3' junctions of the c-*mos*$^{\mathrm{Mo}}$ and Mo-MLV leading to the formation of the v-*mos*$^{\mathrm{Mo}}$ gene, we used DNA fragments from the plasmids p·CS-1, p·CS-2 and p·MLV-1 (Fig. 2). Figure 4 shows the

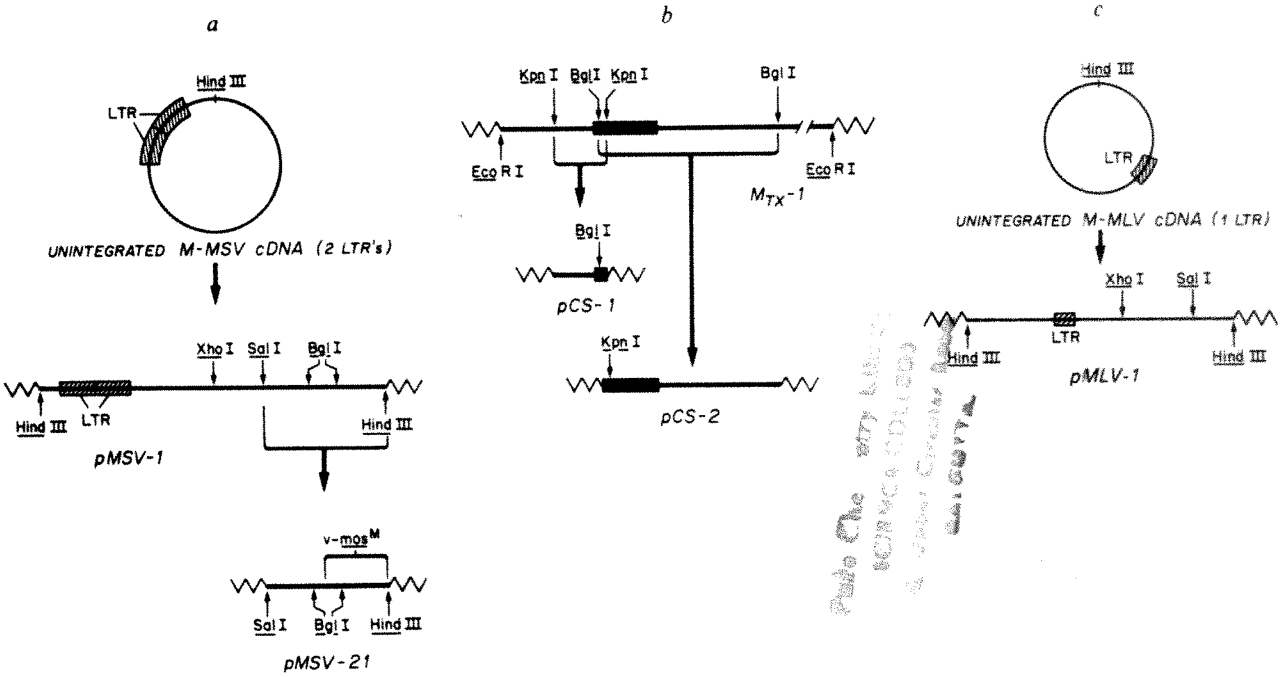
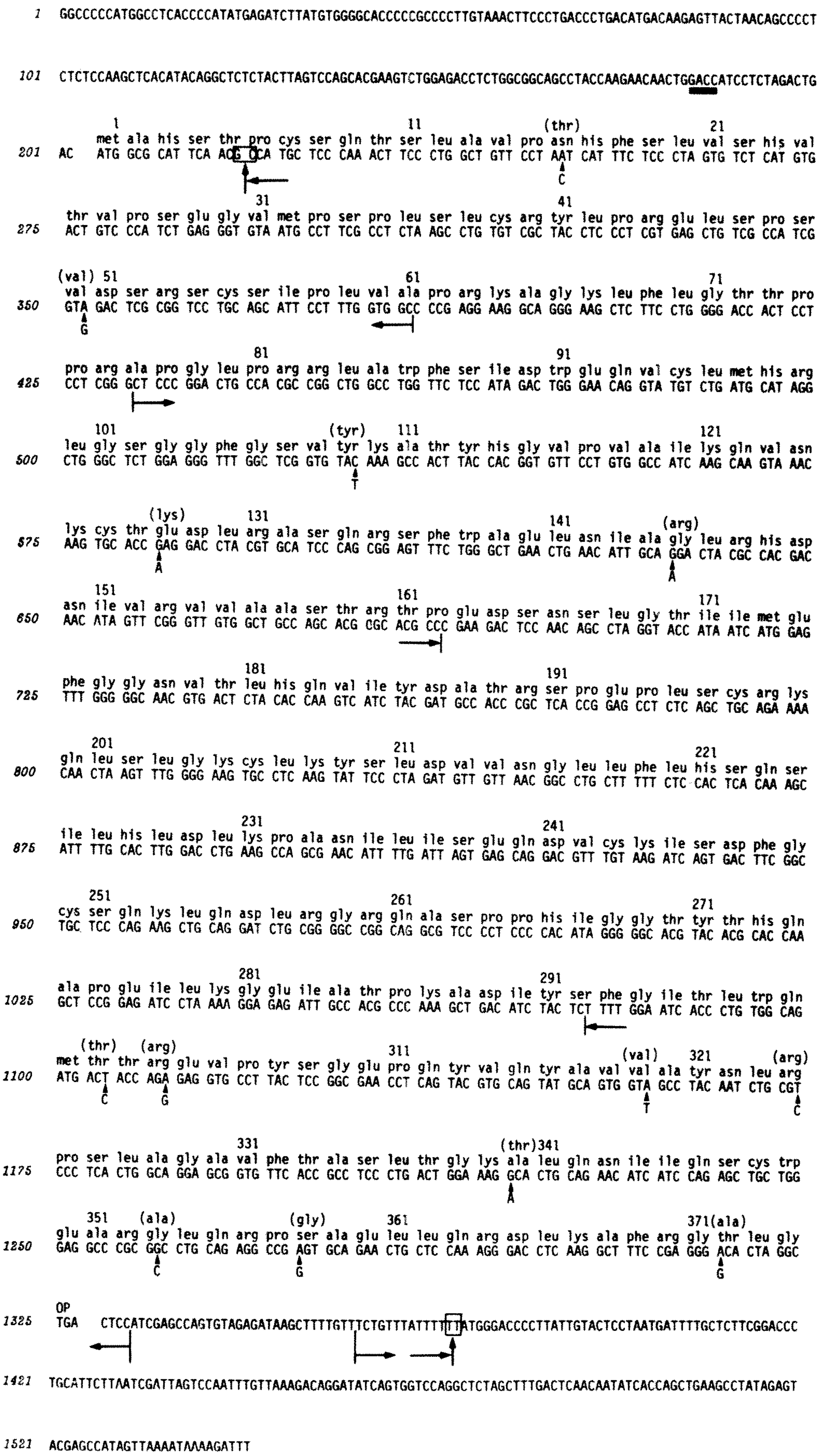


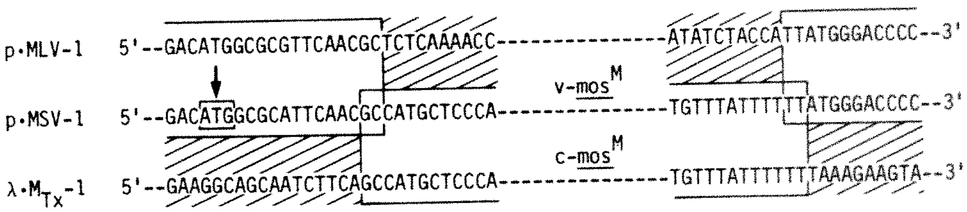
Fig. 2 Construction of plasmids. *a*, p·MSV-1 and p·MSV-21. The unintegrated circular forms of Mo-MSV DNA from freshly infected NIH/3T3 cells were isolated and cleaved at the unique *Hin*dIII site[30]. The DNA was molecularly cloned in the unique *Hin*dIII site of bacteriophage Charon 21A as described elsewhere[25]. The insert DNA containing the entire Mo-MSV DNA with two long terminal repeats was subcloned in the *Hin*dIII site of plasmid pBR322. About 1.0 μg of p·MSV-1 was cleaved with restriction endonucleases *Hin*dIII and *Sal*I and fractionated on 0.7% agarose gels. The 2.2-kilobase *Hin*dIII–*Sal*I fragment containing the v-*mos*$^{\mathrm{Mo}}$ gene was eluted from the gel and joined to *Hin*dIII- and *Sal*I- cleaved pBR322 by T4 DNA ligase as described elsewhere[25]. The recombinant plasmids were used to transform *Escherichia coli* C600SF and colonies were screened both for drug resistance ($Amp^{R}Tet^{S}$) and hybridization to Mo-MSV cDNA probes. The bracket denotes the v-*mos*$^{\mathrm{Mo}}$ sequences. *b*, p·CS-1 and p·CS-2. The 14.0-kilobase *Eco*RI fragment from uninfected NIH/3T3 cells containing the c-*mos*$^{\mathrm{Mo}}$ sequences was cloned in bacteriophage Charon 4A. The recombinant clone, $\lambda \cdot M_{TX}$-1, was analysed by restriction endonuclease[28]. The clone, $\lambda \cdot M_{TX}$-1, was cleaved with *Kpn*I which generated an expected fragment of about 1.2 kilobases, containing the 5' half of c-*mos*$^{\mathrm{Mo}}$ and adjacent cellular sequences. It was eluted from an agarose gel and a stretch of oligo(dC) residues was added using terminal transferase, as described previously[43], and cloned in the *Pst*I-cleaved [dG]-tailed pBR322. To make the plasmid p·CS-2, the recombinant clone, $\lambda \cdot M_{TX}$-1, was cleaved with *Bgl*I and a 5.5-kilobase fragment containing the 3' c-*mos*$^{\mathrm{Mo}}$ and adjacent cellular sequences was isolated from agarose gels and subcloned in the plasmid pBR322 by the [dC]-[dG]-tailing method[43]. The solid box indicates the c-*mos*$^{\mathrm{Mo}}$ sequences. *c*, p·MLV-1. The details of molecular cloning of unintegrated circular forms of Mo-MLV are described elsewhere[26]. Briefly, the Mo-MLV DNA was cleaved at an unique site with *Hin*dIII and cloned in the unique *Hin*dIII site of bacteriophage Charon 21A. The recombinant clone containing the Mo-MLV DNA was cleaved with *Hin*dIII and the insert Mo-MLV DNA further subcloned in the *Hin*dIII site of pBR322. The orientation of the insert in the plasmid was determined by restriction endonuclease analysis. The wavy lines show vector DNA.



```
1      GGCCCCCATGGCCTCACCCCATATGAGATCTTATGTGGGGCACCCCCGCCCCTTGTAAACTTCCCTGACCCTGACATGACAAGAGTTACTAACAGCCCCT

101    CTCTCCAAGCTCACATACAGGCTCTCTACTTAGTCCAGCACGAAGTCTGGAGACCTCTGGCGGCAGCCTACCAAGAACAACTGGACCATCCTCTAGACTG

       1                                       11                 (thr)                21
       met ala his ser thr pro cys ser gln thr ser leu ala val pro asn his phe ser leu val ser his val
201 AC ATG GCG CAT TCA ACG CCA TGC TCC CAA ACT TCC CTG GCT GTT CCT AAT CAT TTC TCC CTA GTG TCT CAT GTG
                         ↑                                          ↑
                                                                    C
                               31                                      41
       thr val pro ser glu gly val met pro ser pro leu ser leu cys arg tyr leu pro arg glu leu ser pro ser
275    ACT GTC CCA TCT GAG GGT GTA ATG CCT TCG CCT CTA AGC CTG TGT CGC TAC CTC CCT CGT GAG CTG TCG CCA TCG

      (val) 51                                      61                                      71
       val asp ser arg ser cys ser ile pro leu val ala pro arg lys ala gly lys leu phe leu gly thr thr pro
350    GTA GAC TCG CGG TCC TGC AGC ATT CCT TTG GTG GCC CCG AGG AAG GCA GGG AAG CTC TTC CTG GGG ACC ACT CCT
         ↑
         G
                               81                                      91
       pro arg ala pro gly leu pro arg arg leu ala trp phe ser ile asp trp glu gln val cys leu met his arg
425    CCT CGG GCT CCC GGA CTG CCA CGC CGG CTG GCC TGG TTC TCC ATA GAC TGG GAA CAG GTA TGT CTG ATG CAT AGG

           101                            (tyr)    111                                     121
       leu gly ser gly gly phe gly ser val tyr lys ala thr tyr his gly val pro val ala ile lys gln val asn
500    CTG GGC TCT GGA GGG TTT GGC TCG GTG TAC AAA GCC ACT TAC CAC GGT GTT CCT GTG GCC ATC AAG CAA GTA AAC
                                             ↑
                                             T
                  (lys)        131                                     141            (arg)
       lys cys thr glu asp leu arg ala ser gln arg ser phe trp ala glu leu asn ile ala gly leu arg his asp
575    AAG TGC ACC GAG GAC CTA CGT GCA TCC CAG CGG AGT TTC TGG GCT GAA CTG AAC ATT GCA GGA CTA CGC CAC GAC
                   ↑                                                                   ↑
                   A                                                                   A
           151                                     161                                     171
       asn ile val arg val val ala ala ser thr arg thr pro glu asp ser asn ser leu gly thr ile ile met glu
650    AAC ATA GTT CGG GTT GTG GCT GCC AGC ACG CGC ACG CCC GAA GAC TCC AAC AGC CTA GGT ACC ATA ATC ATG GAG

                               181                                     191
       phe gly gly asn val thr leu his gln val ile tyr asp ala thr arg ser pro glu pro leu ser cys arg lys
725    TTT GGG GGC AAC GTG ACT CTA CAC CAA GTC ATC TAC GAT GCC ACC CGC TCA CCG GAG CCT CTC AGC TGC AGA AAA

           201                                     211                                     221
       gln leu ser leu gly lys cys leu lys tyr ser leu asp val val asn gly leu leu phe leu his ser gln ser
800    CAA CTA AGT TTG GGG AAG TGC CTC AAG TAT TCC CTA GAT GTT GTT AAC GGC CTG CTT TTT CTC CAC TCA CAA AGC

                               231                                     241
       ile leu his leu asp leu lys pro ala asn ile leu ile ser glu gln asp val cys lys ile ser asp phe gly
875    ATT TTG CAC TTG GAC CTG AAG CCA GCG AAC ATT TTG ATT AGT GAG CAG GAC GTT TGT AAG ATC AGT GAC TTC GGC

           251                                     261                                     271
       cys ser gln lys leu gln asp leu arg gly arg gln ala ser pro pro his ile gly gly thr tyr thr his gln
950    TGC TCC CAG AAG CTG CAG GAT CTG CGG GGC CGG CAG GCG TCC CCT CCC CAC ATA GGG GGC ACG TAC ACG CAC CAA

                               281                                     291
       ala pro glu ile leu lys gly glu ile ala thr pro lys ala asp ile tyr ser phe gly ile thr leu trp gln
1025   GCT CCG GAG ATC CTA AAA GGA GAG ATT GCC ACG CCC AAA GCT GAC ATC TAC TCT TTT GGA ATC ACC CTG TGG CAG

          (thr)   (arg)                            311                            (val)    321        (arg)
       met thr thr arg glu val pro tyr ser gly glu pro gln tyr val gln tyr ala val val ala tyr asn leu arg
1100   ATG ACT ACC AGA GAG GTG CCT TAC TCC GGC GAA CCT CAG TAC GTG CAG TAT GCA GTG GTA GCC TAC AAT CTG CGT
             ↑       ↑                                                               ↑                   ↑
             C       G                                                               T                   C
                               331                            (thr)341
       pro ser leu ala gly ala val phe thr ala ser leu thr gly lys ala leu gln asn ile ile gln ser cys trp
1175   CCC TCA CTG GCA GGA GCG GTG TTC ACC GCC TCC CTG ACT GGA AAG GCA CTG CAG AAC ATC ATC CAG AGC TGC TGG
                                                                   ↑
                                                                   A
           351    (ala)               (gly)        361                                     371(ala)
       glu ala arg gly leu gln arg pro ser ala glu leu leu gln arg asp leu lys ala phe arg gly thr leu gly
1250   GAG GCC CGC GGC CTG CAG AGG CCG AGT GCA GAA CTG CTC CAA AGG GAC CTC AAG GCT TTC CGA GGG ACA CTA GGC
                    ↑                  ↑                                                       ↑
                    C                  G                                                       G
       OP
1325   TGA   CTCCATCGAGCCAGTGTAGAGATAAGCTTTTGTTTCTGTTTATTTTTTATGGGACCCCTTATTGTACTCCTAATGATTTTGCTCTTCGGACCC
                                                            ↑

1421   TGCATTCTTAATCGATTAGTCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGCTTTGACTCAACAATATCACCAGCTGAAGCCTATAGAGT

1521   ACGAGCCATAGTTAAAATAAAAGATTT
```

Fig. 3 Nucleotide and predicted amino acid sequence of v-*mos*$^{\mathrm{Mo}}$ gene. The complete nucleotide sequence of Mo-MSV DNA fragment containing the v-*mos*$^{\mathrm{Mo}}$ gene is shown. The numbers in italics on the left-hand margin refer to the nucleotide numbers (1–1,547). The numbers in the figure (1–374) refer to the predicted amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ putative gene product, starting with Met as number 1 and terminating with an opal codon, TGA. The thick solid arrows under the boxed nucleotides show the recombination junctions between c-*mos*$^{\mathrm{Mo}}$ and Mo-MLV sequences; the horizontal arrows indicate the regions of the c-*mos*$^{\mathrm{Mo}}$ whose nucleotide sequence was determined; the solid triangles and the nucleotides under them denote the base alterations in the c-*mos*$^{\mathrm{Mo}}$ sequences; the amino acids in parentheses above the predicted amino acid sequences of v-*mos*$^{\mathrm{Mo}}$ gene product are the predicted changes in the c-*mos*$^{\mathrm{Mo}}$ product; the closed bar between nucleotides 184 and 187 shows the position of the 192-nucleotide Mo-MLV sequence deleted in Mo-MSV. The locations of some endo·R cleavage sites are: *Acy*I: 986; *Ava*I: 385, 426; *Bal*I: 557; *Bgl*I: 8,448; *Bgl*II: 26; *Cla*I: 1,432; *Hha*I: 208, 681; *Hin*dIII: 1,351; *Hin*fI: 353, 692, 739, 1,083, 1,326, 1,486; *Hpa*I: 844; *Hpa*II: 435, 448, 776, 979, 1,028, 1,125; *Kpn*I: 711; *Pst*I: 368, 794, 966, 1,227, 1,266; *Pvu*II: 789, 1,504; *Sst*II: 1,258; *Xba*I: 192. A complete list of endo·R restriction sites for this sequence is available on request.

Fig. 4 The 5′ and 3′ recombination junctions between Mo-MLV, Mo-MSV and $\lambda \cdot M_{TX}$-1. The 5′ and 3′ boundaries of the recombination junctions between Mo-MLV and c-mos^{Mo} leading to the formation of the v-mos^{Mo} gene sequence are shown. The open boxes denote homology whereas the hatched areas indicate non-homology. The ATG (below the arrow) indicates the position of the predicted first amino acid (Met) of the v-mos^{Mo} gene product (Fig. 3).


p·MLV-1
5'--GACATGGCGCGTTCAACGCTCTCAAAACC--------------ATATCTACCATTATGGGACCCC--3'
v-mos^M
p·MSV-1
5'--GACATGGCGCATTCAACGCCATGCTCCCA--------------TGTTTATTTTTTATGGGACCCC--3'
c-mos^M
λ·M_Tx-1
5'--GAAGGCAGCAATCTTCAGCCATGCTCCCA--------------TGTTTATTTTTTTAAAGAAGTA--3'

sequence of c-mos^{Mo} and p·MLV-1 at the 5′ and 3′ ends of the v-mos^{Mo} gene sequence. The 5′ junction of c-mos^{Mo} with Mo-MLV occurs 12 nucleotides downstream from the initiation codon of the putative v-mos^{Mo} gene product. At the 3′ end of the junction there are 19 nucleotides to the right of the *Hin*dIII site at the end of a stretch of six T residues. The N-terminus of the v-mos^{Mo} gene product extends five amino acids into the Mo-MLV genome past the c-mos^{Mo} junction. There is no direct sequence homology between Mo-MLV and c-mos^{Mo} junction sequences. However, the sequence TTCA is present immediately to the left of the 5′ junction in c-mos^{Mo}, and two nucleotides to the left of the 5′ junction in Mo-MLV (Fig. 4). This kind of sequence homology may suggest that v-mos^{Mo} originated by homologous recombination between c-mos^{Mo} and Mo-MLV sequences. The exact mechanism of recombination, however, remains obscure.

Comparison of portions of c-mos^{Mo} and v-mos^{Mo} sequences

It was of interest to determine the sequence divergence between c-mos^{Mo} and v-mos^{Mo} sequences. A cloned isolate of Mo-MLV (Mo-MLV clone 1) has been obtained[32,33]. We have propagated Mo-MSV clone 124 and Mo-MLV (clone 1) in our laboratory since 1973. Figure 3 compares the nucleotide sequence of portions of v-mos^{Mo} and c-mos^{Mo} sequences. In a stretch of about 500 nucleotides at the 5′ end there seem to be only five base changes. Similarly, in a stretch of 300 nucleotides at the 3′ end there are only eight single base substitutions. The nucleotide sequence homology between the v-mos^{Mo} and c-mos^{Mo} sequences is in complete agreement with the results obtained from nuclease S_1 analysis of hybrids formed between cloned $\lambda \cdot M_{TX}$-1 DNA (containing c-mos^{Mo} sequences, Fig. 2) and Mo-MSV genomic RNA[28]. The initiation codon ATG (at position 203 in Fig. 3) of the longest possible v-mos^{Mo} gene product is not present in the c-mos^{Mo} sequences (Fig. 4). The sequences in c-mos^{Mo} alone are not able to transform cells; however, when portions of Mo-MLV containing the long terminal repeat (LTR) are attached to it, cellular transformation can be observed (W. M. McClements and G. F. van de Woude, personal communication). Thus, it seems that the proteins encoded by v-mos^{Mo} and c-mos^{Mo} sequences are very similar.

The nucleotide sequences adjacent to the 3′ end of the v-mos^{Mo} gene in Mo-MLV and Mo-MSV are identical. However, at the 5′ end, 29 nucleotides from the 5′ junction of c-mos^{Mo} and Mo-MLV, the Mo-MSV seems to have undergone a previously undetected deletion of 192 nucleotides (shown as closed bar in Fig. 3). Analysis of the sequences on the 5′ and 3′ ends of the deletion reveals a common sequence, GACC. This small deletion was not observed in the heteroduplexes formed between Mo-MLV cDNA and Mo-MSV 124 RNA[11,15]. As there are only 29 base pairs between this deletion and the major substitution loop, it was perhaps scored as a part of the major substitution loop, β_S (ref. 11).

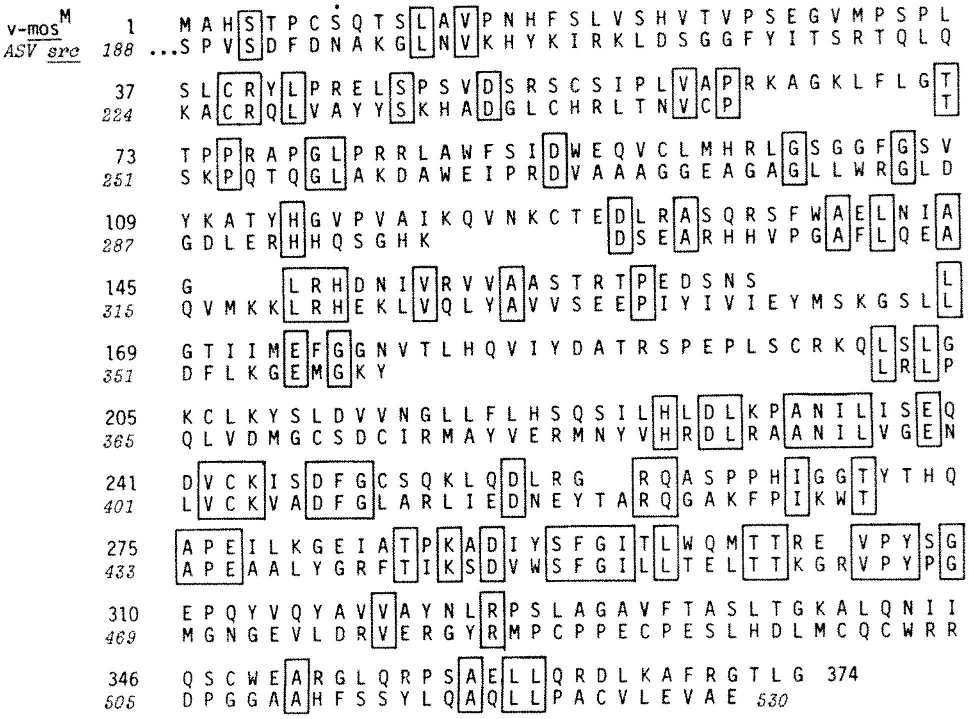

v-mos^M
ASV src
1 M A H S T P C S Q T S L A V P N H F S L V S H V T V P S E G V M P S P L
188 ...S P V S D F D N A K G L N V K H Y K I R K L D S G G F Y I T S R T Q L Q
37 S L C R Y L P R E L S P S V D S R S C S I P L V A P R K A G K L F L G T
224 K A C R Q L V A Y Y S K H A D G L C H R L T N V C P T
73 T P P R A P G L P R R L A W F S I D W E Q V C L M H R L G S G G F G S V
251 S K P Q T Q G L A K D A W E I P R D V A A A G G E A G A G L L W R G L D
109 Y K A T Y H G V P V A I K Q V N K C T E D L R A S Q R S F W A E L N I A
287 G D L E R H H Q S G H K D S E A R H H V P G A F L Q E A
145 G L R H D N I V R V V A A S T R T P E D S N S L
315 Q V M K K L R H E K L V Q L Y A V V S E E P I Y I V I E Y M S K G S L L
169 G T I I M E F G G N V T L H Q V I Y D A T R S P E P L S C R K Q L S L G
351 D F L K G E M G K Y L R L P
205 K C L K Y S L D V V N G L L F L H S Q S I L H L D L K P A N I L I S E Q
365 Q L V D M G C S D C I R M A Y V E R M N Y V H R D L R A A N I L V G E N
241 D V C K I S D F G C S Q K L Q D L R G R Q A S P P H I G G T Y T H Q
401 L V C K V A D F G L A R L I E D N E Y T A R Q G A K F P I K W T
275 A P E I L K G E I A T P K A D I Y S F G I T L W Q M T T R E V P Y S G
433 A P E A A L Y G R F T I K S D V W S F G I L L T E L T T K G R V P Y P G
310 E P Q Y V Q Y A V V A Y N L R P S L A G A V F T A S L T G K A L Q N I I
469 M G N G E V L D R V E R G Y R M P C P P E C P E S L H D L M C Q C W R R
346 Q S C W E A R G L Q R P S A E L L Q R D L K A F R G T L G 374
505 D P G G A A H F S S Y L Q A Q L L P A C V L E V A E 530

Fig. 5 Comparison between the proposed amino acid sequences for the v-mos^{Mo} and ASV *src* gene products. The amino acid sequences of the v-mos^{Mo} (Fig. 3) and ASV *src* (ref. 36) gene products were aligned for optimum identity as described elsewhere[37,38]. The portions of the sequences presented here include the entire v-mos^{Mo} sequence, and ASV *src* sequence from amino acid 188 to its carboxy terminus (residue 530). Areas of identity are depicted by boxes. The letter code for the amino acids is: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

The sequence in Fig. 3 was examined for features which have been implicated in RNA synthesis in other systems. Neither the v-*mos*$^{\mathrm{Mo}}$ nor the c-*mos*$^{\mathrm{Mo}}$ sequences seem to contain the canonical sequence (TATAAA, Hogness–Goldberg box) often implicated in promotion of transcription. Examination of the A+T-rich region of the sequence shown in Fig. 3 reveals a sequence, TTGTAAA, beginning at nucleotide 53, which might be used as a DNA-dependent RNA polymerase II binding site. Further sequence analysis of Mo-MLV, 5′ from nucleotide 1, reveals another potential promoter sequence (data not shown). The importance of the TATAAA sequence is uncertain, however, as it has recently been reported that removal of TATAAA sequences does not eliminate the transcription of histone genes *in vivo*[34]. To the 3′ side of the v-*mos*$^{\mathrm{Mo}}$ coding region, ~210 nucleotides from the TGA stop codon, there is the sequence AATAAA which has been associated with the addition of poly(A) tails onto eukaryotic mRNAs[35].

Comparison with other proteins

The amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product, as derived from the DNA sequence, was subjected to a computer search for sequence resemblance among ~1,300 known amino acid sequences, including those for the T-antigens of polyoma and SV40. Although several sequences were found to have some marginal resemblances to portions of the v-*mos*$^{\mathrm{Mo}}$ gene product, only the recently reported oncogene product (deduced from DNA sequence studies) of avian sarcoma virus (ASV)[36] had a genuinely homologous sequence.

A very significant homology was uncovered when overlapping segments of the two proposed amino acid sequences were compared (Fig. 5). Residues 1–374 of the v-*mos*$^{\mathrm{Mo}}$ gene product were placed into optimum alignment with residues 188–530 (C-terminus) of the ASV *src* gene product using a program which has an algorithm similar to that described by Needleman and Wunsch[37,38], and which uses an empirically derived gap penalty. After an alignment score had been obtained, the two sequences were randomly jumbled, 36 comparisons of jumbled sequences executed and the alignment scores again computed. A mean value of the alignment scores was, finally, determined and the standard deviation calculated. The alignment score obtained for the two authentic sequences was about 7 s.d. above the mean of the jumbled comparisons, a highly significant result[38].

The optimum alignment of the two sequences revealed that they are about 23% identical in the region of the 371-residue overlap shown in Fig. 5. The strongest homology occurs over the stretch of the v-*mos*$^{\mathrm{Mo}}$ sequence corresponding to residues 227–309 and the ASV sequence residues 387–468. In this region there are 36 identical residues over the course of the 83-residue span, or about 45% identity. The overall degree of resemblance is approximately the same as that obtained in a comparison of mouse haemoglobin α-chain and hagfish haemoglobin[39], or of the β-chain of human haemoglobin[40] and human myoglobin[41]. The murine and avian haemoglobin α-chains, by contrast, have evolved less rapidly than have the proposed murine and avian viral proteins: the mouse haemoglobin α-chain has about 70% identity with chicken α-haemoglobin.

No detectable nucleotide sequence homology was observed when ASV *src*-specific cDNA was hybridized to viral RNA obtained from the Mo-MSV(MLV) complex[42]. This result was confirmed by a computer comparison of the nucleotide sequences of v-*mos*$^{\mathrm{Mo}}$ (Fig. 3) and the ASV *src* gene product[36]. However, the amino acid sequence homology between the oncogene products of ASV and Mo-MSV indicates that there can be little doubt that the two proteins have evolved from a common ancestor.

Conclusion

We have studied the biogenesis of the transforming gene of Mo-MSV. The Mo-MSV was isolated from sarcomas obtained after passage of Mo-MLV in mice. The Mo-MLV acquired cellular sequences, and underwent several deletions to yield a transforming but replication-defective virus. We have determined the entire nucleotide sequence of the Mo-MSV-specific sequences, referred to as the v-*mos*$^{\mathrm{Mo}}$ gene, in the Mo-MSV. Furthermore, the recombination junctions between the Mo-MLV and the cellular sequences leading to the formation of the v-*mos*$^{\mathrm{Mo}}$ gene have also been determined. The v-*mos*$^{\mathrm{Mo}}$ and c-*mos*$^{\mathrm{Mo}}$ share an uninterrupted stretch of 1,159 nucleotides with few base changes. There are no apparent structural homologies between the Mo-MLV and c-*mos*$^{\mathrm{Mo}}$ sequences at the point of recombination. However, there seems to be a small sequence homology, TTCA (Fig. 4), at the 5′ junction of Mo-MLV and c-*mos*$^{\mathrm{Mo}}$, perhaps suggesting homologous recombination. Due to the absence of specific tumour antisera, no authentic v-*mos*$^{\mathrm{Mo}}$ gene product has been identified. The 37K *in vitro* translation product[31] is specifically inhibited if the RNA is hybridized to a v-*mos*$^{\mathrm{Mo}}$ gene DNA fragment before translation (J. Papkoff, T. Hunter and I. M. V., unpublished results). The predicted amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product can now be used to verify the nature of the 37K protein synthesized *in vitro*. Finally, the proposed amino acid sequence of the v-*mos*$^{\mathrm{Mo}}$ gene product has been shown to share significant homology with the ASV *src* gene product.

We thank Bart Sefton and Walter Eckhart for reading the manuscript and Carolyn Goller for preparation of it, Drs P. Czernilofsky and Hugo Martinez for help with computer programs, Dr E. P. Reddy for communicating his sequencing data on Mo-MSV^{src}-specific fragment before publication, and M. Jones and D. Murdock for help with various stages of the study. This work was supported by grants from the NCI and the American Cancer Society. D.J.D. acknowledges a Helen Hay Whitney Foundation postdoctoral fellowship.

Received 21 July; accepted 11 November 1980.

1. Bishop, J. M. *A. Rev. Biochem.* **47,** 37 (1978).
2. Spector, D. H., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4102 (1978).
3. Duesberg, P., Bister, K. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4320 (1977).
4. Hayman, M. J., Roy-Pokora, B. & Graf, T. *Virology* **92,** 31 (1979).
5. Rettenmier, C. W., Anderson, S. M., Rieman, M. W. & Hanafusa, H. *J. Virol.* **32,** 749 (1979).
6. Lee, W.-H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 2018 (1980).
7. Hanafusa, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 3009 (1980).
8. Reynolds, F. H. Jr, van de Ven, W. J. M. & Stephenson, J. R. *J. Virol.* **28,** 665 (1978).
9. Witte, O. N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2488 (1978).
10. van de Ven, W. J. M., Kahn, A. S., Reynolds, F. H. Jr, Mason, K. T. & Stephenson, J. R. *J. Virol.* **33,** 1034 (1980).
11. Hu, S., Davidson, N. & Verma, I. *Cell* **10,** 469 (1977).
12. Chien, Y.-H. *et al. J. Virol.* **31,** 752 (1979).
13. Hu, S. F., Lai, M. M. C. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1265 (1979).
14. Shields, A., Goff, S., Paskind, M., Otto, G. & Baltimore, D. *Cell* **18,** 955 (1979).
15. Donoghue, D. J., Sharpe, P. A. & Weinberg, R. A. *J. Virol.* **32,** 1015 (1979).
16. Sherr, C. J., Fedele, F. A., Oskarsson, M., Maizel, J. & Van de Woude, G. *J. Virol.* **34,** 200 (1980).
17. Van de Woude, G. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 4464 (1979).
18. Mellon, P., Pawson, A., Bister, K., Martin, G. S. & Duesberg, P. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5874 (1977).
19. Van Zaane, D. & Bloemers, H. P. J. *Biochim. biophys. Acta* **516,** 249 (1978).
20. Aaronson, S. A. & Rowe, W. P. *Virology* **42,** 9 (1970).
21. Scolnick, E. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **72,** 4650 (1975).
22. Moloney, J. B. *Natn. Cancer Inst. Monogr.* **22,** 139 (1966).
23. Ball, J. K., McCarter, J. A. & Sunderland, S. M. *Virology* **56,** 268 (1973).
24. Dina, D., Beemon, K. & Duesberg, P. *Cell* **9,** 299 (1976).
25. Verma, I. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 1773 (1980).
26. Berns, A. *et al. J. Virol.* (in the press).
27. Oskarsson, M., McClements, W. M., Blair, D. G., Maizel, J. V. & van de Woude, G. F. *Science* **207,** 1222 (1980).
28. Jones, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 2651 (1980).
29. Benz, E. D. Jr & Dina, D. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3294 (1979).
30. Verma, I. M. *Nucleic Acids Res.* **6,** 1863 (1979).
31. Papkoff, J., Hunter, T. & Beemon, K. *Virology* **101,** 91 (1980).
32. Moloney, J. B. *J. natn. Cancer Inst.* **24,** 933 (1969).
33. Fan, H. & Paskind, M. *J. Virol.* **14,** 421 (1974).
34. Grosschedl, H. & Birnstiel, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1432 (1980).
35. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263,** 211 (1976).
36. Czernilofsky, A. P. *et al. Nature* **287,** 198 (1980).
37. Needleman, S. B. & Wunsch, C. D. *J. molec. Biol.* **48,** 443 (1970).
38. Jue, R. A., Woodbury, N. W. & Doolittle, R. F. *J. molec. Evol.* **15,** 129 (1980).
39. Liljeqvist, G., Braunitzer, G. & Paleus, S. *Hoppe-Seyler's Z. physiol. Chem.* **360,** 125 (1979).
40. Braunitzer, G. *et al. Hoppe-Seyler's Z. physiol. Chem.* **325,** 283 (1961).
41. Herrera, A. E. R. & Lehmann, M. *Nature new Biol.* **232,** 149 (1971).
42. Stehelin, D., Guntaka, R. V., Varmus, H. E. & Bishop, J. M. *J. molec. Biol.* **101,** 349 (1976).
43. Van Beveren, C., Goddard, J. G., Berns, A. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3307 (1980).

A model from electron microscopy for the molecular structure of fibrinogen and fibrin

John W. Weisel, George N. Phillips Jr & Carolyn Cohen

Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254

The overall architecture of the fibrinogen molecule and aspects of its packing to form fibrin have been derived from a study of electron microscope images of ordered arrays. The molecule, of length 450 Å, is seen in more detail than the Hall–Slayter model and is made up of seven globular domains connected by rod-like segments.

FIBRINOGEN is the soluble precursor of the fibrin clot in the plasma. On tissue injury, an enzymatic clotting cascade is initiated leading to thrombin cleavage of the two small pairs of highly charged peptides from fibrinogen[1,2]. The molecule then becomes highly insoluble and rapidly aggregates into the fibrin clot. This noncovalently bonded polymer is cross-linked by the fibrin-stabilizing factor enzyme and is later dissolved by cleavage with plasmin. The processes of clotting and fibrinolysis both occur continually in the normal functioning of the vascular system. The complexity of this clotting process is reflected in the complex structure of the fibrinogen molecule.

Fibrinogen has a molecular weight of 340,000 and an elongated form[3,4] but its detailed shape is unknown. Hall and Slayter[5] obtained electron micrographs of individual molecules, revealing a protein about 475 Å long consisting of a linear array of three globular domains (Fig. 1*A*). Despite considerable controversy, this 'trinodular' model is now generally accepted[5–11].

To account for the small 223 Å axial period of fibrin, Hall and Slayter invoked a 'shrinkage' of the molecule, but their proposal is inconsistent with results of X-ray diffraction studies, which show no large-scale changes in molecular structure on the conversion of fibrinogen to fibrin[12,13]. The axial period of fibrin can be accounted for most simply by a half-staggered arrangement of fibrinogen monomers[13].

Chemical studies indicate that fibrinogen is a dimer: the molecule consists of three pairs of different polypeptide chains (called Aα, Bβ and γ) linked by 29 disulphide bridges[14,15]. The complete amino acid sequence of human fibrinogen and the complex pattern of disulphide cross-links have recently been established[16–22]. Sequence analysis has also been used to locate certain regions of α-helical coiled coil[23,24] that had been predicted to connect the globular domains of the molecule[25]. However, these studies do not provide direct evidence on the overall morphology of the molecule or its packing in fibrin.

The most powerful approach to a determination of the structure of fibrinogen is X-ray crystallographic analysis. An important discovery was that microcrystals can be produced on limited cleavage of bovine fibrinogen with a bacterial protease[26] and that with further digestion, large single crystals, ordered to about 4 Å resolution, can be grown[27]. Crystallographic and chemical evidence show that the modified molecule in the crystal is only 5–10% smaller in mass than native fibrinogen[28]. Moreover, the molecules forming the crystals clot and the fibrin so produced has a band pattern in the electron microscope very similar to that of native fibrin, indicating that the modified protein is largely intact (Fig. 2*D*, *E*). (Other proteolytic enzymes have also been shown to modify fibrinogen so that microcrystals and crystals can be grown[29].)

Our strategy for determining the low resolution crystallographic structure of fibrinogen is the refinement of a model against X-ray data[30]. Here we derive a more detailed molecular model for fibrinogen than the trinodular Hall–Slayter structure and describe packing arrangements that account for electron microscope images in a variety of highly ordered arrays including fibrin.

Development of the model

It has been recognized for some time that the electron microscope band pattern of fibrin cannot be accounted for by a simple Hall–Slayter model. Within the repeat of 225 Å in negative contrast there is one bright strain-excluding band about 60 Å wide and three narrow light bands, the middle one being somewhat less prominent or variable in intensity[31–34] (Fig. 2*D*). The size and number of bands in this image show quite plainly that the globular domains in the fibrinogen molecule must be considerably smaller than those of the Hall–Slayter model (Fig. 1*A*) and that a trinodular structure cannot give rise to the details of the staining pattern.

A key to the interpretation of the fibrin band pattern can be found in the so-called 'orthogonal sheet' structure[26,28] (Fig. 2*A*). A simple, half-staggered arrangement of this highly regular array gives a band pattern closely resembling that of negatively stained fibrin[28]. An important aspect of the orthogonal sheet form is its extremely detailed substructure. The features are such that neither the Hall–Slayter model, nor, indeed, any simple arrangement of molecules running parallel to the microcrystal axis can readily account for this image. However, analysis of a variety of microcrystals of fibrinogen, including the orthogonal

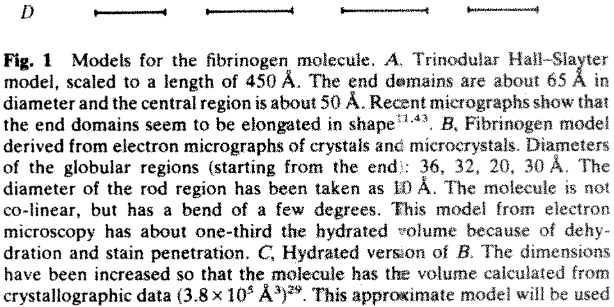


Fig. 1 Models for the fibrinogen molecule. *A*, Trinodular Hall–Slayter model, scaled to a length of 450 Å. The end domains are about 65 Å in diameter and the central region is about 50 Å. Recent micrographs show that the end domains seem to be elongated in shape[11,43]. *B*, Fibrinogen model derived from electron micrographs of crystals and microcrystals. Diameters of the globular regions (starting from the end): 36, 32, 20, 30 Å. The diameter of the rod region has been taken as 10 Å. The molecule is not co-linear, but has a bend of a few degrees. This model from electron microscopy has about one-third the hydrated volume because of dehydration and stain penetration. *C*, Hydrated version of *B*. The dimensions have been increased so that the molecule has the volume calculated from crystallographic data (3.8×10^5 Å^3)[29]. This approximate model will be used in the initial phasing of the X-ray data. *D*, Predicted α-helical coiled-coil regions based on analysis of the amino acid sequence[23,24]. The length of the region between two 'braces' of disulphide bonds is 111 residues, or about 160 Å; the coiled coil would begin at the central domain. We infer that the smallest globular domain may correspond to a region predicted to be non-helical where plasmin cleavage occurs[23] (see text).

0028–0836/81/040263—05$01.00

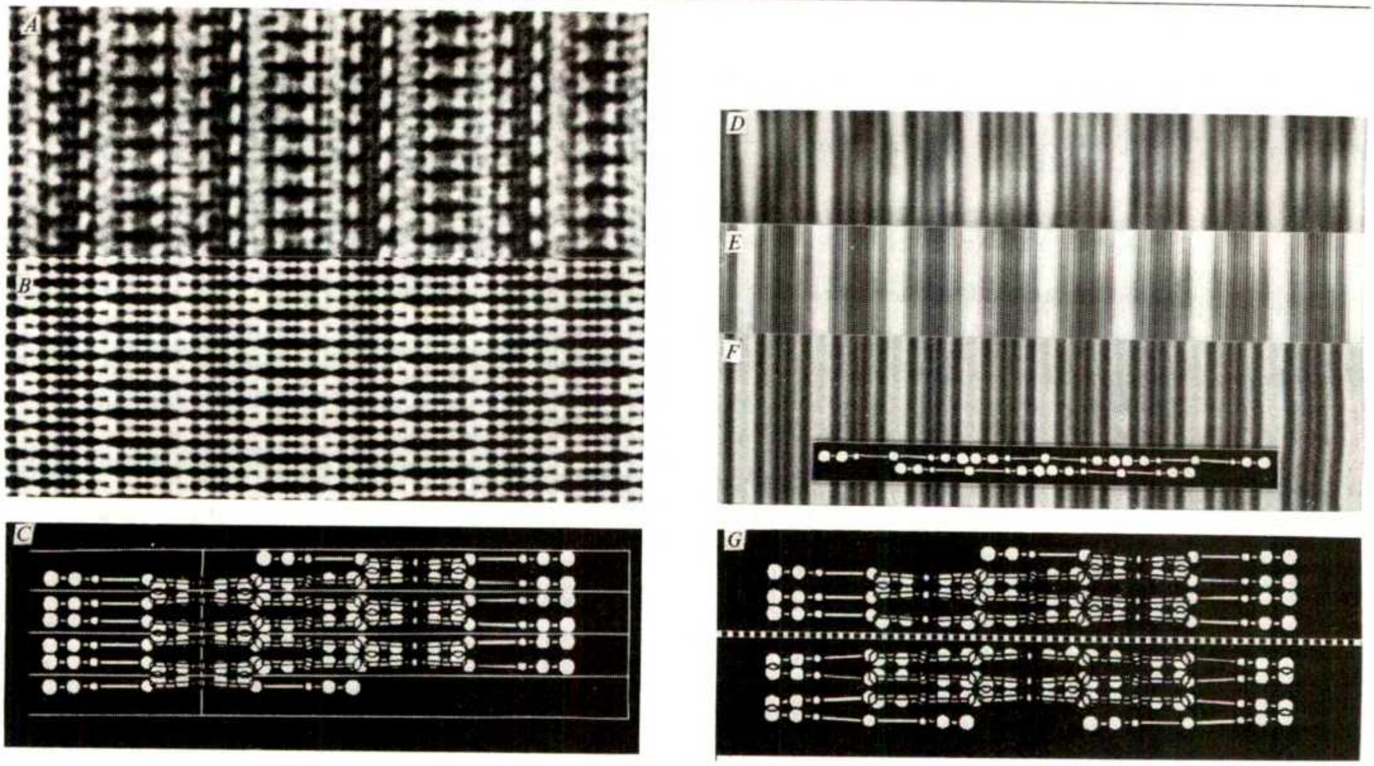


Fig. 2 Simulation of electron microscope images of the orthogonal sheet microcrystals and fibrin. *A*, Electron micrograph of a negatively contrasted microcrystal of the orthogonal sheet form[28]; the image has been optically filtered to remove background noise[44,45]. Unit cell dimensions are 900 × 90 Å. ×450,000. *B*, TV display of computer simulation of *A*, based on the model below (*C*) (as described in the text). Note that the models in *C* and *G* are drawn to a scale about 1.5 times as great as the electron micrographs above. *D*, *E*, Computer filtered electron micrographs of negatively contrasted fibrin. The axial repeat is 225 Å. In *D*, the fibrin is derived from native fibrinogen whereas in *E* the fibrin is from *Pseudomonas*-modified fibrinogen. Note that the weakest light band, which arises from the smallest globular domain, is less prominent in the modified fibrin[26,28], suggesting that the bacterial protease, like plasmin (see text), cleaves some material from this region. *F*, TV display of computer simulation of fibrin, based on the model shown in *G*. The fibrin band pattern was generated by superposition of arrays of the orthogonal sheet, shown in *A*, shifted by 225 Å (refs 26, 28). For clarity, two such layers of molecules are shown, separated by a dashed line, in the upper and lower portions of *G*. The lateral contacts in both fibrin and the orthogonal sheet form incorporate a half-stagger, where the central globular domain of one molecule fits into the 'pocket' formed by the end-to-end linkage of two molecules in an adjacent filament. A simple half-staggered array of five molecules is shown as an inset to *F* (and with the same scale) so that the relationship of the molecular shape and packing to the fibrin band pattern is apparent. The bright axial striations in both the orthogonal sheet form and fibrin arise from this juxtaposition of the end globular units of two molecules with the middle unit of a third. In the orthogonal sheet, the pairs of small bright units on either side of this feature can be seen to arise from the adjacent globular domains in these molecules. To account for the angle of the thin strands in the gap, the molecules are arrayed with a slight tilt. This arrangement generates the correct symmetry for the structure and also places the smallest globular units at the centre of the dark gap where they are seen as bright nodes.

sheet form, has indicated that all these structures seem to be made up of 450 Å-long fibrinogen molecules that bond end to end[28].

The best characterized macrocrystal form of fibrinogen is the monoclinic $P2_1$ crystal of protease-modified bovine fibrinogen[27,28]. A packing model for this crystal was established using the morphology and special features of the diffraction patterns[29,35]. Here too it was shown that the molecules are 450 Å long and bond end to end to form filaments. The general run of the molecules in the lattice was established but not their lateral relations. This knowledge of the molecular packing allowed an interpretation of electron microscope images of crystal fragments. In particular, a characteristic net or mesh form was observed that could be accounted for as the view along the unique symmetry axis (Fig. 3*A*, *D*). The strands of the mesh seen in the electron microscope were interpreted as pairs of symmetry-related molecular filaments in a distorted cell. These strands showed prominent features of macromolecular size, but again it was not possible using the Hall–Slayter model to define the lateral packing or to delineate individual molecules.

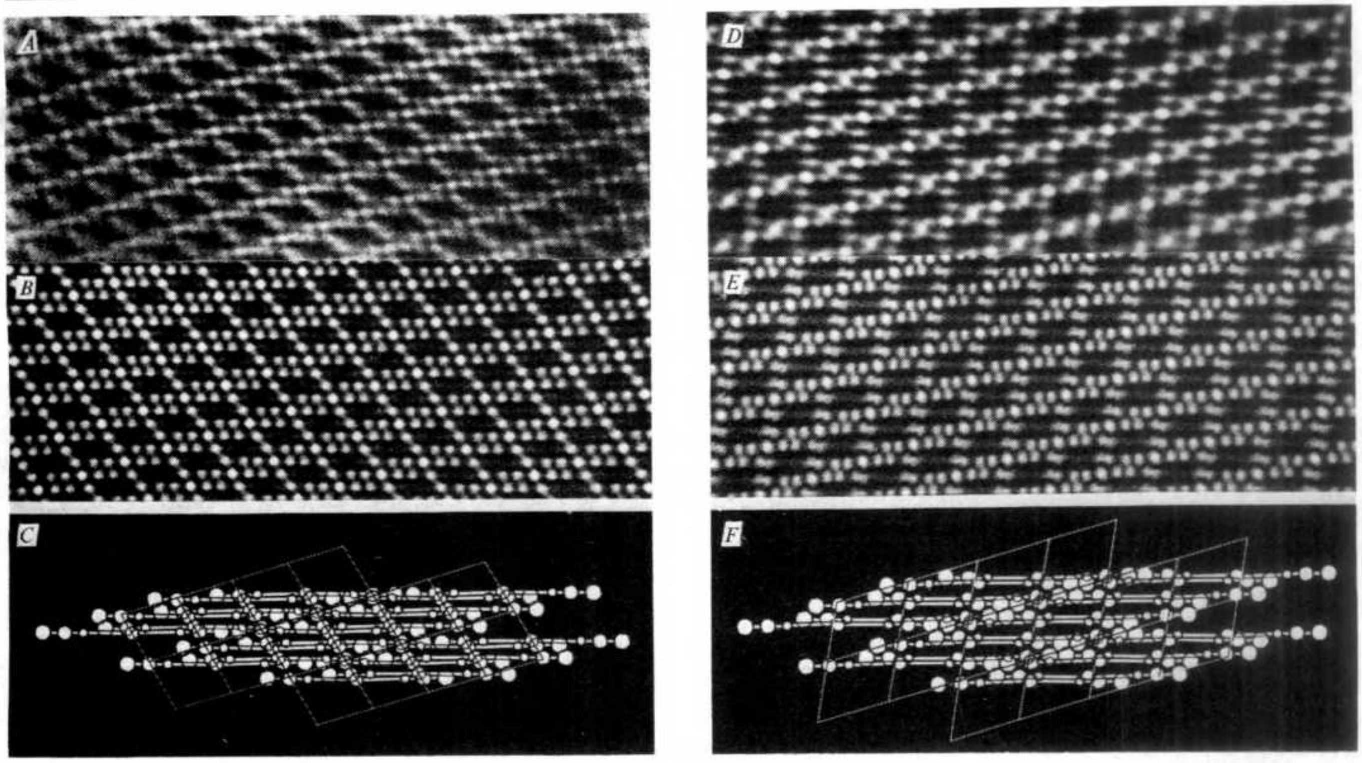
A form that we found useful in establishing the model was a 'segment' aggregate obtained from human fibrinogen modified by bacterial protease[36] (Fig. 4). The segments displayed a strikingly simple band pattern in negative staining that revealed where the molecules begin and end. Analysis of this structure showed that the molecules are in register in layers which are tilted alternately with respect to the long axis. Here again, the molecules were found to be ~450 Å long.

Because there is no molecular overlap, the packing of fibrinogen in this form yields direct information about the mass distribution along the molecule. The overall structure seems qualitatively like the Hall–Slayter model in that there are three major stain-excluding domains per molecular length. There is, however, additional fine structure in the segments. Within each major band there is another thin line of opposite contrast, which seems to define smaller domains. This evidence again suggested that the Hall–Slayter model was valid as a very low resolution image of the molecule, but that a more detailed mass distribution could be derived.

We may interpret the staining of the segments to mean that the outer domain of the Hall–Slayter model is, in fact, made up of two units, each about 30–40 Å in diameter. Close inspection of the segment images reveals an additional thin white line within the broad dark-staining region (Fig. 4). This feature indicates that another smaller globular unit may be located between the outer domain and the central nodule. This additional globular unit is, in fact, necessary to account for the detailed appearance of images of both fibrin and related forms as well as views of the crystal (see below). The molecule of length 450 Å thus appears in all the ordered aggregates as a heptad made up of seven globular domains. We show below how more precise dimensions of the globular units, as well as those of the rod-like regions connecting them, can be derived from the simulation of a variety of electron microscope images.

Simulation of electron microscope images

Using modifications of the model described above, computer simulations were carried out to develop a single molecular model whose substructure and packing could account for the negatively stained images of a variety of arrays seen in the



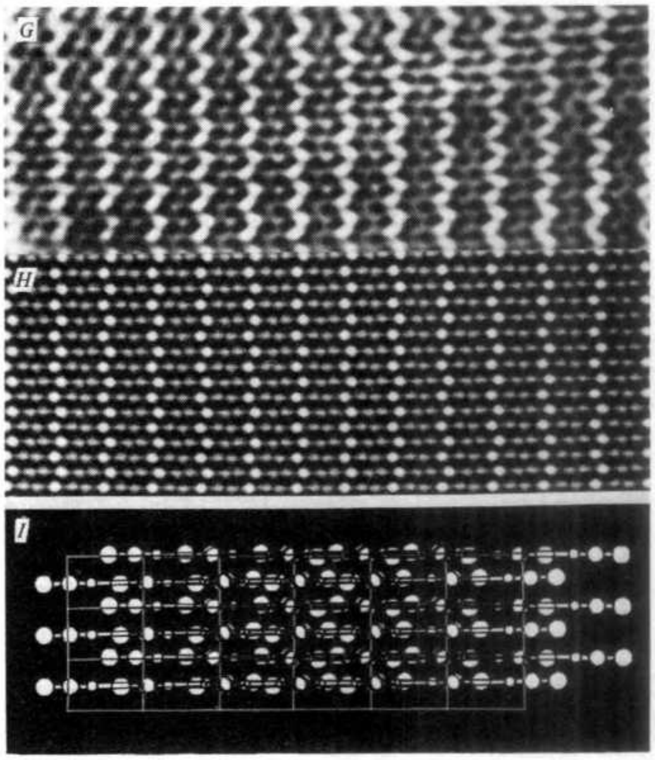


Fig. 3 Simulation of electron microscope images of the $P2_1$ crystal. *A*, *D*, *G*, Electron micrographs of negatively contrasted crystal fragments which have been filtered optically[44,45] or by computer processing of a digitized image[45,46]. *A* and *D* are two different mesh images corresponding to crystals with different amounts of shrinkage viewed along the unique axis. In both cases, the molecular strands are oriented parallel to the largest face of the crystal, which is horizontal here[29]. Unit cell parameters in *A* are: $a = 132$, $b = 97$, $c = 125$, $\beta = 105°$, whereas in *B* they are: $a = 160$, $b = 97$, $c = 130$, $\beta = 65°$. The image in *G* corresponds to the state of the crystal seen in *D* but viewed in the perpendicular direction. ×450,000. *B*, *E*, *H*, TV display of computer simulations of the models in *C*, *F*, *I* below (as described in the text). Note that the models in *C*, *F* and *I* are drawn to a scale about 1.5 times as great as the electron micrographs and simulations above so as to display the molecular packing. In the crystal, the strands correspond to the projection of filaments in different layers staggered by roughly one-third of the molecular length. In both the mesh images (*A*, *D*) the relative brightness of the different units of the motif depends on the size of the globular domains and their degree of overlap. A cluster of three overlapping globular domains corresponds to the broad light region nearest the dark gap; the very bright unit at the middle of the group corresponds to two overlapping globular regions. A small bend accounts in part for the slightly wavy appearance of the strands. The rod portion of the molecule can be visualized in some micrographs of the mesh, as in *D*.

electron microscope. Trial models for selected forms were rapidly generated with the aid of computer graphics and compared with electron micrographs. The molecule was represented as a series of seven spheres connected by rods, and the programs were written so that one set of parameters fixed the sizes and position of the spheres and rods within one molecule, and another set located a single molecule in the unit cell. Additional molecules were then generated according to the appropriate symmetry elements of the specific form under analysis. To convert these arrangements to density contours, which could then be compared with electron microscope images, the model was represented by sets of points spaced closely enough to correspond to a continuous structure at this resolution. The projected Fourier transform of this configuration was then computed, followed by the inverse transform which was displayed on a television screen. Subjective criteria were used to alter both molecular and packing parameters in comparing the simulation to filtered electron microscope images.

A range of molecular models was examined. Slight differences in the size of the globular units or slight departures of the two halves of the molecule from co-linearity were tested to achieve a good fit with a particular form. Some differences in the nature of the end-to-end molecular contacts were also tried with certain arrays. It was found that all the images could be fitted with the same heptad model (Fig. 1*B*).

The $P2_1$ crystal

Several images from $P2_1$ crystal fragments could be simulated. The net described previously and identified as the view down the unique axis of the crystal is characterized by strands having features of macromolecular size running in the horizontal direction[29] (Fig. 3*A*). The repeating unit along the strands is a group of seven stain-excluding regions of varying sizes (none larger than ~40 Å)—the middle one being particularly bright—followed by a dark-staining gap. This motif repeats every 450 Å and corresponds to the projection of two symmetry-related layers of molecular filaments in the unit cell. This form was particularly useful in establishing the size of the globular domains and their separation in the model. The small bend in the molecule accounts in part for the slightly wavy appearance of the strands. With a shift of about one-third the molecular length between filaments in the two layers the simulation shows excellent agreement with the filtered image of the electron micrograph (Fig. 3*B*, *C*).

Electron micrographs of the $P2_1$ crystal fragments show a number of states of the mesh described above. Another commonly observed form has a mesh with a different appearance (Fig. 3*D*). The motif is rather similar to that described above, although the intensities and positions of the bright units differ somewhat. Keeping the parameters of the molecular model used in the previous simulation, only a slight change in the relative stagger of the molecular filaments within each layer is required for a satisfactory fit (Fig. 3*E*, *F*).

A different view of the crystal could also be identified. Some preparations contained very thin crystals that tended to lie with the large face on the grid. This image showed a distinctive pattern of alternating rows of bright and weak zig-zags repeated 145 Å axially (Fig. 3*G*). The spacings indicated that this was the view along one of the crystal axes perpendicular to the previous view and nearly perpendicular to the large face of the crystal (see Fig. 3*D*). Using the same arrangement of molecules as before (Fig. 3*F*), and simply taking the appropriate projection (Fig. 3*I*), a very satisfactory simulation can be achieved (Fig. 3*H*).

The orthogonal sheet form and fibrin

One of the most detailed images is that of the orthogonal sheet form, related to fibrin (Fig. 2*A*). Here the motif is a very complex dumbell-shaped bright region about 400 Å long with a dark-staining gap of about 50 Å. The motifs in adjacent rows are staggered by half the lateral period so that the true axial repeat is 900 Å. Within the dark-staining gap, thin light strands seem to connect units in different rows. For the simulation, the dimensions of the molecular model as determined for the $P2_1$ crystal views were maintained, but the molecular filaments in different layers were half-staggered, as determined previously[13,26,28] (Fig. 2*C*). This arrangement accounts for the very bright axial striation which arises from juxtaposition of the end globular units of two molecules with the middle unit of a third (Fig. 2*B*). The small bright units on either side of this feature can be seen to arise from the adjacent globular domains in these molecules. To account for the angle of the thin strands in the gap, the molecules were arrayed with a slight tilt to the long axis of the unit cell. This arrangement generates the correct symmetry for the structure. The smallest globular units can be seen quite clearly as bright nodes at the centre of the dark gaps.

A half-stagger (225 Å) of this arrangement for the orthogonal sheet form then generates the fibrin image[26,28] (Fig. 2*E*, *F*, *G*). The brightest band in fibrin again arises from juxtaposition of the end globular regions of two molecules together with a centre domain of a neighbouring molecule and the narrower band on either side from the adjacent domains. Additional features of the pattern are also accounted for: the thinnest band in the period corresponds to the smallest globular domain in the molecule. In fact, this arrangement is precisely that proposed in the early half-staggered models for fibrin[13], but the size of the domains in the molecular model now yields correct dimensions for the band pattern (Fig. 2*F* inset). The relatively simple band pattern of fibrin is thus accounted for by a superposition of rather complex molecular arrays.

The molecular model

We have now obtained an improved low-resolution model for the fibrinogen molecule (Fig. 1*B*). The basic structure is closely related to the trinodular Hall–Slayter model, but we have been able to establish a precise length for the molecule of 450 Å, and a more detailed substructure. The outer globular domain, incompletely resolved by the shadowing technique used by Hall and Slayter, is now seen to consist of two regions of nearly equal size. The narrow rod-like connection between the central and outer units had not previously been visualized, and we have shown that its regular run is interrupted by a small globular region displaced towards the outer domains. There is also a slight bending of the molecule by a few degrees. We consider the evidence in favour of this model to be compelling: all the features of the model are directly visible in certain electron microscope images, and the model can account for the detailed appearance of a variety of forms, including the complex pattern of fibrin.

Fibrinogen displays different properties at different stages of clotting. As shown for other fibrous proteins[37], particular domains may be associated with each of its various functions. The central and outer globular domains of the molecule are involved in assembly of the fibrin fibre; the middle region contains the small, highly charged fibrinopeptides cleaved by thrombin[38] whereas the outer globular domains contain complementary bonding sites for fibrin formation[39,40]. Covalent cross-linking of the molecules by the fibrin-stabilizing factor has been shown to involve the outer globular domain[41]. Some of these sites may also be located in the neighbouring globular domain. The plasmin-sensitive portion of the molecule can be identified with the smallest globular domain which interrupts the regular run of the rod-like α-helical coiled-coil regions (Fig. 1*B*, *D*)[23,24]. On the basis of this oversimplified picture, we can begin to localize a number of the characteristic properties of fibrinogen in different structural domains.

These results also indicate that a modest degree of molecular 'flexing' is a necessary feature of the packing arrangements in all the forms we have examined. Hence, the molecule should not be considered as rigid, but there is no compelling reason to postulate marked departures from colinearity of the two halves of the molecule.

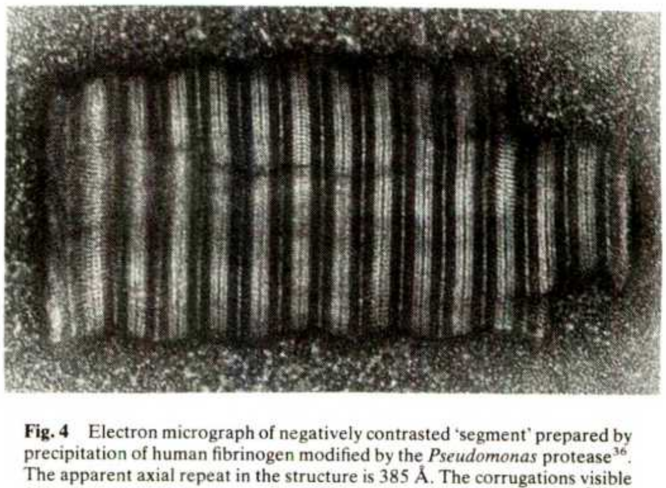


Fig. 4 Electron micrograph of negatively contrasted 'segment' prepared by precipitation of human fibrinogen modified by the *Pseudomonas* protease[36]. The apparent axial repeat in the structure is 385 Å. The corrugations visible along the upper and lower edges indicate that the true period is 770 Å and that the molecules are both tilted in layers and bent. The end of the segment on the right shows where the molecules begin. ×135,000.

The model described here represents the first stage in our approach to the detailed three-dimensional structure of fibrinogen. The dimensions used in our simulations were derived from negatively stained electron microscope images, but the approximate hydrated envelope can be calculated from the molecular volume in the crystal[29] (Fig. 1*C*). We have now completed three-dimensional X-ray data collection to about 20 Å on the $P2_1$ crystal form. To derive truer shapes of the molecular domains we are refining the hydrated version of the model against the diffraction data[42]; solvent flattening techniques may also be useful for phasing reflections at this resolution[30]. The new model will reveal in even greater detail the complex architecture of this highly evolved protein molecule.

We thank Cynthia Stauffacher and Nancy Tooney for helpful discussions, Paul Norton for electron microscopy, William Saunders for photography, David DeRosier for computer programs and Edward Swinarski for computing. This work was supported by NIH grant AM17346 to C.C. and NSF grant PCM78-15913 to D.D. J.W.W. is an Established Investigator of the American Heart Association; G.N.P. held a National Research Service Award from the USPHS and is a Medical Foundation Fellow.

Received 4 July; accepted 4 November 1980.

1. Davie, E. W. & Fujikawa, T. *A. Rev. Biochem.* **44,** 799–829 (1975).
2. McFarlane, R. G. in *Human Blood Coagulation, Haemostasis and Thrombosis* (ed. Biggs, R) 1–31 (Blackwell, Oxford, 1976).
3. Shulman, S. *J. Am. chem. Soc.* **75,** 5846–5852 (1953).
4. Caspary, E. A. & Kekwick, R. A. *Biochem. J.* **67,** 41–48 (1957).
5. Hall, C. E. & Slayter, H. S. *J. biophys. biochem. Cytol.* **5,** 11–15 (1959).
6. Köppel, G. *Nature* **212,** 1608–1609 (1966).
7. Krakow, W., Endres, G. F., Siegel, B. M. & Sheraga, H. A. *J. molec. Biol.* **71,** 95–103 (1972).
8. Bachmann, L., Schmitt-Fumian, W. W., Hammel, R. & Lederer, K. *Makromolec. Chem.* **176,** 2603–2618 (1975).
9. Marguerie, G. & Stuhrmann, H. B. *J. molec. Biol.* **102,** 143–156 (1976).
10. Slayter, H. S. *Ultramicroscopy* **1,** 341–357 (1976).
11. Fowler, W. E. & Erickson, H. P. *J. molec. Biol.* **134,** 241–249 (1979).
12. Bailey, K., Astbury, W. T. & Rudall, K. M. *Nature* **151,** 716–717 (1943).
13. Stryer, L., Cohen, C. & Langridge, R. *Nature* **197,** 793–794 (1963).
14. Henschen, A. *Ark. Kemi* **22,** 355–373 (1964).
15. Henschen, A. *Thromb. Haemostasis* **42,** 14 (1979).
16. Blombäck, B., Hessel, B. & Hogg, D. *Thromb. Res.* **8,** 639–658 (1976).
17. Henschen, A. & Lottspeich, F. *Hoppe-Seyler's Z. physiol. Chem.* **357,** 1801–1803 (1976).
18. Doolittle, R. F. *Horizons Biochem. Biophys.* **3,** 164–191 (1977).
19. Lottspeich, F. & Henschen, A. *Hoppe-Seyler's Z. physiol. Chem.* **358,** 935–938 (1977).
20. Watt, K. W. K., Tagaki, T. & Doolittle, R. F. *Biochemistry* **18,** 68–76 (1979).
21. Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D. & Riley, M. *Nature* **280,** 464–468 (1979).
22. Henschen, A., Lottspeich, F. & Hessel, B. *Hoppe-Seyler's Z. physiol. Chem.* **360,** 1951–1956 (1979).
23. Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. *J. molec. Biol.* **120,** 311–325 (1978).
24. Parry, D. A. D. *J. molec. Biol.* **120,** 545–551 (1978).
25. Cohen, C. *J. Polymer Sci.* **49,** 144–145 (1961).
26. Tooney, N. M. & Cohen, C. *Nature* **237,** 23–25 (1972).
27. Cohen, C. & Tooney, N. M. *Nature* **251,** 659–660 (1974).
28. Tooney, N. M. & Cohen, C. *J. molec. Biol.* **110,** 363–385 (1977).
29. Weisel, J. W., Warren, S. G. & Cohen, C. *J. molec. Biol.* **126,** 159–183 (1978).
30. Phillips, G. N., Lattman, E. E., Cummins, P., Lee, K. Y. & Cohen, C. *Nature* **278,** 413–417 (1979).
31. Cohen, C., Revel, J. P. & Kucera, J. *Science* **141,** 436–438 ((1963).
32. Cohen, C., Slayter, H. S., Goldstein, L., Kucera, J. & Hall, C. *J. molec. Biol.* **22,** 385–388 (1966).
33. Kay, D. & Cuddigan, B. J. *Br. J. Haemat.* **13,** 341–347 (1967).
34. Karges, H. E. & Kuhn, K. *Eur. J. Biochem.* **14,** 94–97 (1970).
35. Tooney, N. M. & Weisel, J. W. in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* Vol. 1 (eds Parry, D. A. D. & Creamer, L. K.) 393–427 (Academic, London, 1979).
36. Weisel, J. W., Tooney, N. M., Kaplan, I., Amrani, D. & Cohen, C. *J. molec. Biol.* **143,** 329–334 (1980).
37. Cohen, C. in *Molecular Architecture in Cell Physiology* (eds Hayashi, T. & Szent-Györgyi, A. G. 169–190 (Prentice-Hall, New Jersey, 1966).
38. Blombäck, B. *et al. Nature* **218,** 130–134 (1968).
39. Blombäck, B., Hessel, B., Hogg, D. & Therkildsen, L. *Nature* **275,** 501–505 (1978).
40. Olexa, S. A. & Budzynski, A. Z. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1374–1378 (1980).
41. Fowler, W. E., Erickson, H. P., Hantgan, R. R., McDonagh, J. & Hermans, J. *Science* (in the press).
42. Stauffacher, C., Phillips, G. N., Weisel, J. W., Ford, E. D. & Cohen, C. (in preparation).
43. Slayter, H. S. in *Principles and Techniques of Electron Microscopy* Vol. 9 (ed. Hayat, M.) 175–245 (Van Nostrand, New York, 1978).
44. Klug, A. & DeRosier, D. J. *Nature* **212,** 29–32 (1966).
45. Erickson, H. P., Voter, W. A. & Leonard, K. *Meth. Enzym.* **49,** 39–63 (1978).
46. Misell, D. L. in *Practical Methods in Electron Microscopy* Vol. 7 (ed. Glauert, A. M.) (North-Holland, Amsterdam, 1978).

LETTERS

Cosmic ray antiprotons and modified closed galaxy model

S. A. Stephens

Tata Institute of Fundamental Research, Homi Bhabha Road,
Bombay 400005, India

Observations suggest that a small flux of secondary antiprotons ($\bar{p}$), created in the interaction of cosmic ray nuclei with interstellar gas, should also be present in the primary cosmic radiation. Calculations based on accelerator data for the production of $\bar{p}$ and existing models for the propagation of cosmic rays, predict too small a flux of $\bar{p}$ compared with the finite flux observed recently[1]. Although the observed excess could be attributed to the existence of primary cosmic ray antimatter, the experimental upper limits[2,3] of the ratio of antinuclei to nuclei of charge $|Z| \geqslant 2$ seem to contradict this. Here we use a 'modified closed galaxy' model to attribute this excess to secondary antiprotons. In this model, ~50% of the observed cosmic ray nucleons are of recent origin and they propagate according to the 'nested leaky box' model, while the rest propagate according to 'closed galaxy' model. This model explains the observations on $\bar{p}$ and e^+, and predicts more D and ^{3}He than do the existing models.

The secondary $\bar{p}$ spectrum in cosmic rays has been calculated in the past by Gaisser and Maurer[4], and Badhwar *et al.*[5] using the extensive accelerator data available for the production of $\bar{p}$. However, these calculated spectra differ by more than a factor of two at ~10 GeV and by even larger factors at lower energies, suggesting that cross-sectional representations used by them must have been responsible for this difference. The interpretation of a finite flux of $\bar{p}$ depends crucially on the accuracy of the estimated flux of $\bar{p}$ in cosmic rays. Szabelski *et al.*[6] showed that the results of Badhwar *et al.*[5] are in error. Thereby, they questioned the 'standard leaky box' model (SLBM) for the propagation of cosmic rays. I recently obtained a more accurate parametrization[7] of the invariant cross-sections, for the production of $\bar{p}$ in inclusive reactions, and calculated the production spectrum of $\bar{p}$. This calculation confirmed the conclusions of Szabelski *et al.*[6]. The results from this calculation[7] are now used to examine the implication of the observed $\bar{p}$ in terms of cosmic ray propagation models.

The equilibrium spectrum of $\bar{p}$ has been obtained by taking into account ionization loss, annihilation of $\bar{p}$, and inelastic interactions in which $\bar{p}$ retains nearly half its energy. The resultant $\bar{p}/p$ ratio as a function of the kinetic energy of $\bar{p}$ is shown in Fig. 1 for various cosmic ray models. Curve S is the estimated $\bar{p}/p$ ratio according to SLBM, in which cosmic rays traverse 5 g cm^{-2} of interstellar gas independent of energy. This curve can be compared with other existing calculations[4–6] carried out using SLBM only. The calculations of Gaisser and Maurer[4], and Szabelski *et al.*[6] agree within 25% with curve S over the entire energy region above 2 GeV. This small deviation arises partly from the differences in the parametrization of cross-sectional data and partly due to different demodulated proton spectra used. Badhwar *et al.*[5] overestimate the $\bar{p}/p$ ratio by a factor of 2 at ~10 GeV and this difference increases with decreasing energy; however, they also agree within 25% at energies $\geqslant$50 GeV.

The SLBM has been abandoned since it was observed that the abundance ratio of secondary nuclei (like Li, Be and B) to primary nuclei (like C and O) decreased with increasing energy. This result suggested that high-energy cosmic rays traverse less amount of matter. Recent observations on the spectra of cosmic ray nuclei suggest[8] that the energy dependence of the matter traversal is approximately proportional to $(8\,\mathrm{GV}/c/R)^{1/2}$, where R is the rigidity in GV/c, for $R \geqslant 8$ GV/c and is nearly constant for $R < 8$ GV/c. This dependence has been explained in two ways. (1) The confinement of cosmic rays in the Galaxy is energy dependent[9]—'modified leaky box' model (MLBM). (2) A certain fraction, f, of the matter traversed by cosmic rays in the source regions is energy dependent—'nested leaky box' model[10] (NLBM). Curve M in Fig. 1 is the calculated $\bar{p}/p$ ratio on the basis of MLBM and the hatched region (N) is according to NLBM. The energy-dependent escape probability in NLBM from the source regions is assumed to be similar to that of MLBM, and that the median energy of interacting nucleons is ~10 times the energy of $\bar{p}$ produced; the upper and lower bounds are for $f = 0.5$ and 0.8 respectively. Figure 1 shows that

NLBM predicts too small a flux of $\bar{p}$ compared with MLBM in the energy region below a few tens of GeV.

Figure 1 plots the observed $\bar{p}/p$ ratio of $(5.2\pm1.5)\times10^{-4}$ reported by Golden *et al.*[1] in the energy region 4.7–11.6 GeV; this result was based on 28 $\bar{p}$ events of extraterrestrial origin. The correct value may be higher by ~15% because the expected flux[7] of $\bar{p}$ of atmospheric origin is smaller than that estimated by Golden *et al.* The $\bar{p}/p$ ratio obtained by Bogomolov *et al.*[11] is not plotted because it was based only on two $\bar{p}$ events, and in this experiment, splash albedo particles in the atmosphere could not be eliminated[12]. Figure 1 shows that popular models of cosmic ray propagation predict too low a flux of $\bar{p}$ by a factor ≥4 compared with the observed flux; however, the significance of this difference is only about 3σ. In the following, we examine the 'closed galaxy' model originally proposed by Rasmussen and Peters[13] to understand the apparently high flux of $\bar{p}$ observed by Golden *et al.*[1].

In the 'closed galaxy' model, cosmic rays are totally confined until they are destroyed by interaction. Curve C in Fig. 1 is the calculated $\bar{p}/p$ ratio on the basis of total confinement. The observed ratio falls short of curve C by a factor of ~2. It is therefore proposed that ~50% of the observed cosmic ray nucleons are of recent (~10^7 yr) origin, and are propagated according to either MLBM or NLBM; the remaining ones are the result of total confinement. We can now combine 50% of curve C with either curve M or hatched area N, and the resultant $\bar{p}/p$ ratio lies between the two short curves passing through the data point in Fig. 1. Thus, the 'modified closed galaxy' model (MCGM) can account for the observed flux of $\bar{p}$.

The secondary positron spectrum in cosmic rays can now be calculated using MCGM. The equilibrium spectrum of e^+ is calculated by using the recently estimated[14] production spectrum and taking into account all energy-loss processes—ionization, bremsstrahlung, synchrotron and inverse Compton scattering. For convenience, the values of mean magnetic field $\mathbf{B}_\perp$ and the radiation density are assumed to be the same for all models and are taken to be 4 μG and 0.7 eV cm^{-3} respectively.

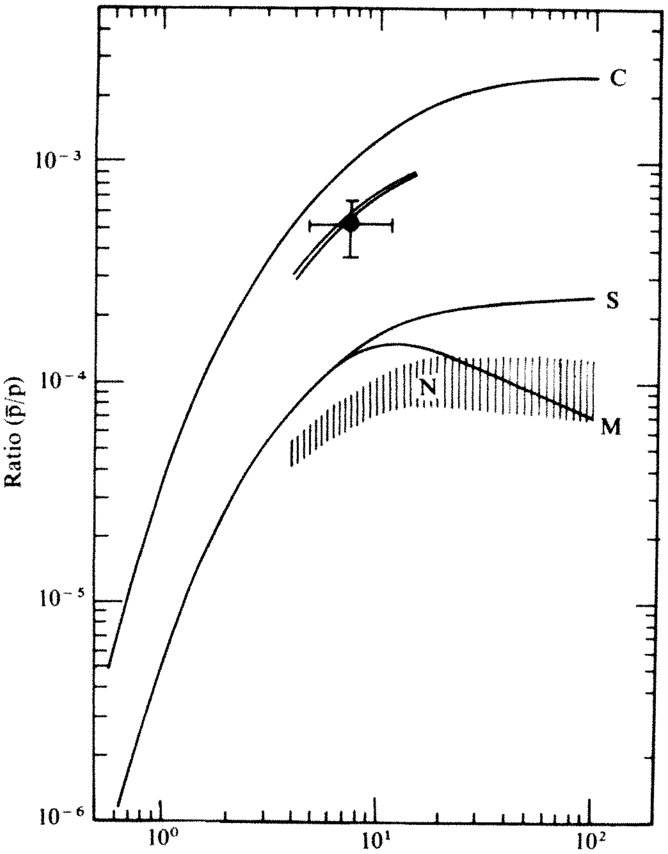


Fig. 1 The calculated $\bar{p}/p$ ratio plotted as a function of the kinetic energy of $\bar{p}$, for various models of cosmic ray propagation. The solid curves passing through the data point are the estimated ratio by summing 50% of curve C with 50% of either curve M or the lower bound of the hatched area N. ●, Data from ref. 1; S, SLBM; M, MLBM; N, NLBM; C, CGM.

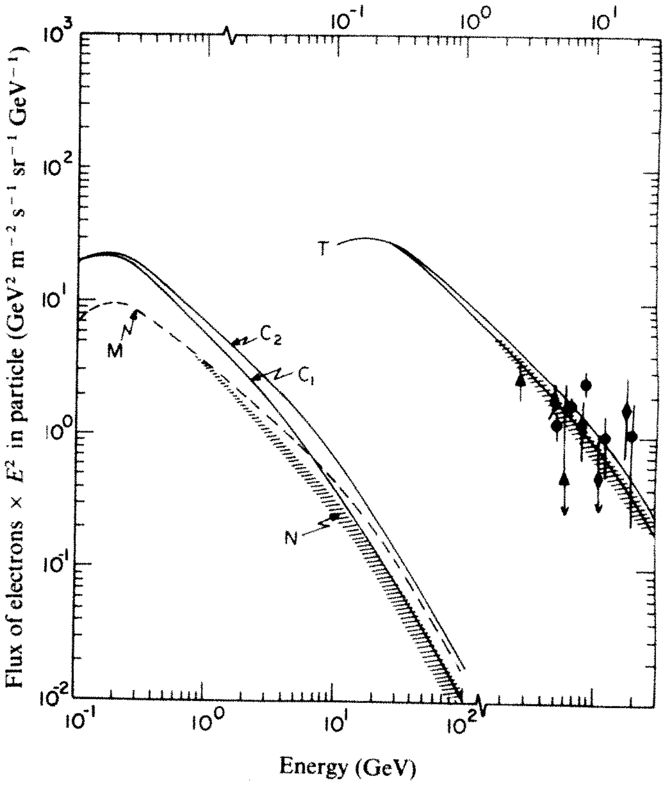


Fig. 2 The calculated energy spectra of the old and new components of secondary cosmic ray e^+ are shown in the left side, and the combined e^+ and total electron spectra on the right side; the flux values are multiplied by E^2. Symbols as in Fig. 1; curves C_1 and C_2 correspond to $n_H=0.1$ and 0.2 atom cm^{-3}, the mean hydrogen density of the interstellar gas traversed by the old component. Data are taken from: ◆, ref. 17; ●, ref. 18; ▲, ref. 19.

The calculated spectra of e^+ from the old component are shown in Fig. 2 by curves C_1 and C_2, which correspond respectively to $n_H=0.1$ and 0.2 atom cm^{-3}, where n_H is the mean hydrogen density of the interstellar gas traversed by this component. The flux values in Fig. 2 are multiplied by E^2. The flux of e^+ is proportional to n_H at high energies, where synchrotron and inverse Compton scattering losses dominate, and is independent of n_H at low energies, where ionization and bremsstrahlung losses dominate[15]. The dashed curve M is the calculated new component due to MLBM. In this calculation, the mean lifetime of cosmic rays of $R\leq8$ GV/c is assumed to be 10^7 yr, as determined by radioisotope measurements[16], and during this period they traverse 5 g cm^{-2} of interstellar gas. The recent e^+ component according to NLBM is calculated as in the case of $\bar{p}$ and is shown by the hatched area N.

The sum of the old and new components of the e^+ spectrum is shown on the right side in Fig. 2. The solid curves (T) are the sum of curves M and either C_1 or C_2, and the hatched area corresponds to the sum of the hatched area N with curve C_1 or C_2. The observed flux values of e^+ are plotted as filled data points at energies ≥3 GeV (refs 17–19), where the effect due to solar modulation is small. The band representing MCGM agrees with the observations, which have a large scatter. Note that other existing models cannot satisfactorily explain e^+ observations[6,12]. Also the value of n_H needed for the old component cannot be very much smaller than 0.1 atom cm^{-3}. This means that the effective confinement region for cosmic rays may not include regions very far away from the galactic disk, say larger than a few kpc. Also the total (e^++e^-) spectrum in interstellar space derived on the basis of curve T can be shown to be consistent with that deduced from the non-thermal radio spectrum in the Galaxy.

Half of the observed nucleons, belonging to the old component, is shared among various nuclear species approximately in proportion to the respective interaction mean-free paths. The estimated percentages of the old component among the observed p, He, (C+O), (F to Sc), (P to Cr) and (Fe Group) nuclei are respectively 58, 21, 9, 6, 4 and 3. Note that the new component constitutes $>90\%$ of the observed flux for nuclei other than p and He. Therefore, one can explain the observed energy dependence of the abundance ratio of secondary to primary nuclei using MLBM or NLBM for the new component. Further, the relative abundances of heavy nuclei of charge ≥ 6 would not differ from other models by more than a few per cent.

We now need to choose between MLBM and NLBM for the new component. According to MLBM, the energy spectrum of nucleons of recent origin above ~ 8 GV/c (~ 4 GeV/n) would be steeper than the injection spectrum by a power-law spectral index of about 0.5; the old component retains the injection spectrum. This would mean that the old component would dominate the observed flux beyond a certain energy[20]. The calculated energies at which the intensities from these two components become comparable are ~56, 460 and 1,950 GeV/n for He, (C+O) and (Fe group) respectively; protons have 58% old component even below 8 GeV. This clearly indicates that the observed similarity in the spectral shape of p and He above a few GeV/n cannot be understood, unless protons are accelerated differently. One should also expect a flattening of He spectrum beyond about 50 GeV/n, which has not been observed at least up to a few hundred GeV/n. These two inconsistencies discredit MLBM, and the natural choice is for NLBM, because the spectral shape of the new component is the same as the old.

A reliable estimate of the abundance of deuterium can be made using this model at relativistic energies. The estimated D/He ratio from this model is nearly twice as much from other models at energies ≤ 4 GeV/n and decreases slightly at high energies due to the energy-dependent confinement of the recent component in source regions. A similar enhancement is expected in the case of ^{3}He abundance. The available measurements have been restricted to energies ≤ 200 MeV/n and these results differ by a large factor[21]. However, they can be easily detected at relativistic energies using magnetic spectrometry[22]. The observation of an enhanced abundance of D would not only confirm the predictions of this model but would also reduce the difficulty in understanding the observed charge ratio of muons at sea level. Perhaps the distribution and spectral shape of high-energy γ rays, and the extended nonthermal radio emission in the Galaxy might also be explained by this model.

I thank Professors M. V. K. Apparao, R. R. Daniel and P. V. Ramana Murthy for helpful criticisms.

Received 15 September, accepted 10 November 1980

1 Golden, R L *et al* *Phys Rev Lett* **43**, 1196–1199 (1979)

2 Badhwar, G D *et al* *Nature* **274**, 137–139 (1978)

3 Smoot, G F, Buffington, A & Orth, C D *Phys Rev Lett* **35**, 258–261 (1975)

4. Gaisser, T K & Maurer, R H *Phys Rev Lett* **30**, 1264–1267 (1973)

5 Badhwar, G D, Golden, R L, Brown, M L & Lacy, J L *Astrophys Space Sci* **37**, 283–300 (1975)

6 Szabelski, J, Wdowczyk, J & Wolfendale, A W *Nature* **285**, 386–388 (1980)

7 Stephens, S A *Astrophys Space Sci* (in the press)

8 Orth, C D, Buffington, A, Smoot, G F & Mast, T S *Astrophys J.* **226**, 1147–1161 (1978)

9. Juliusson, E, Meyer, P & Muller, D *Phys Rev Lett* **29**, 445–448 (1972).

10 Cowsik, R & Wilson, L W *Proc 13th int Cosmic Ray Conf* **1**, 500–505 (1973)

11 Bogomolov, E A *et al* *Proc 16th int Cosmic Ray Conf* **1**, 330–335 (1979)

12 Daniel, R R & Stephens, S A *Bull astr Soc Ind* (in the press)

13 Rasmussen, I L & Peters, B *Nature* **258**, 412–413 (1975)

14 Badhwar, G D & Stephens, S A *Proc 15th int Cosmic Ray Conf* **1**, 398–403 (1977)

15 Badhwar, G D & Stephens, S. A. *Phys Rev* D **14**, 356–358 (1976)

16 Buffington, A, Orth, C D. & Mast, T S *Astrophys J* **226**, 355–371 (1978)

17 Fanselow, J L, Hartman, R C, Hildebrand, R H & Meyer, P *Astrophys J* **158**, 771–780 (1969)

18 Golden, R L *et al* Preprint PSL-PE00947 (1979)

19 Buffington, A & Orth, C D *Astrophys J* **199**, 669–679 (1975)

20 Peters, B & Westergaard, N J. *Astrophys Space Sci* **48**, 21–46 (1977)

21 Daniel, R R & Stephens, S A. *Space Science Review* **17**, 45–158 (1975)

22 Badhwar, G D, Stephens, S A & Golden, R L *Proc 14th int Cosmic Ray Conf* **9**, 3183–3187 (1975)

Expanding haloes in cometary comae

W.-H. Ip

Max-Planck-Institut fur Aeronomie, D-3411 Katlenburg-Lindau 3, FRG

Observations indicate that cometary activity can be rather violent. Flare-ups, sunward-pointing jets and periodic formation of concentric haloes expanding outwards can also be persistent features in some comets. For example, a halo structure was observed in the coma of Comet Halley during its last visit[1]. It is argued here that halo formation may be related to the propagation of shock pairs driven through the cometary atmosphere as a result of spin-modulation of the cometary outgassing rate. Even though the non-stationary process could be less dramatic than the example considered here, its effect on the chemistry and expansion of the neutral gas as well as dynamics of the cometary plasma must be significant.

Although theoretical models of cometary coma generally assume that neutral atmospheres expand spherically, symmetrically and without time variation, inhomogeneous structures have often been observed in the cometary comae. These include the formation of concentric shells (or haloes) moving outwards from the central source in succession and the appearance of fans or jets. Such non-stationary behaviour has long been thought to be related to the spin-modulation of the outgassing rate from the cometary nucleus[2,3].

There are two possiblities. If the surface composition of the cometary nucleus is uniform, the dependence of the surface temperature on the angle of insolation (θ) would mean higher gas emission rate at the point shifted from the sunlit position with $\theta = 0°$. (The exact position of the 'hot spot' is determined by the direction and period of the nuclear spin as well as thermal properties of the surface material; the resultant jet effect has been used to explain the nongravitational acceleration or deceleration of comets[2].) As the jetting of the volatile gas—though localized—is more or less constant, the expansion of the cometary atmosphere, except the persistent sunward jets, should show no marked time variation. On the other hand, if the surface of the cometary nucleus can be divided into one active hemisphere and an inactive one (say), there will be more violent ejection of the volatile gas every time the active region turns towards the Sun. In the extreme case, there will be regular outbursts at the centre with a period equal to that of the spin period leading to large spatial and temporal variations in the structure of the coma. Clearly, the formation of expanding haloes as seen in the case of Comet Donati in 1858 (ref. 4) and Comet Halley[1] must be somehow related to such periodic outbursts. (From measurements of their respective halo structures, Whipple[5,6] has determined the spin period of Comet Donati to be 4.6 h and for Comet Halley to be 10.3 h.)

To understand this phenomenon, time-dependent expansion of the gas outflow must be considered. Here it is assumed that the cometary atmosphere is an ideal gas and that the radial expansion is spherically symmetric. Furthermore, the time variation of the outgassing rate is approximated by a step function and we are basically interested in the initial propagation of the cometary neutral gas as the strong source is switched on. If ρ is the density, u the radial velocity, p the pressure and $\gamma(=5/3)$ the ratio of specific heats of the gas, then

$$\frac{\partial \rho}{\partial t}+\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\rho u)=0 \tag{1}$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial t}=-\frac{1}{\rho}\frac{\partial p}{\partial r} \tag{2}$$

$$\frac{\partial}{\partial t}\left(\frac{p}{\rho^\gamma}\right)+u\frac{\partial}{\partial r}\left(\frac{p}{\rho^\gamma}\right)=0 \tag{3}$$

0028-0836/81/040269—03$01 00

for the hydrodynamic equations of expansion of the neutral gas from a point source. Introducing the similarity variable[7],

$$\eta = r/\alpha t \tag{4}$$

for the expansion of a piston gas with a constant speed α and defining

$$u = rV(\eta)/t, \quad \rho = \rho_0 r_0^2 \sigma(\eta)/r^2, \quad p = \rho_0 r_0^2 \pi(\eta)/t^2 \tag{5}$$

with the density distribution of the cometary coma at $t \leq 0$ given as:

$$\rho = \rho_0 r_0^2 / r^2 \tag{6}$$

(that is, $\rho = \rho_0$ at $r = r_0$ for the unperturbed flow), we can find an equivalent set of equations for V, σ and π with η as the sole variable. As discussed by Simon and Axford[8], similarity solutions of V, σ and π can be obtained for the motion of such a piston gas describing a pair of shocks (S_+ and S_-) separated by a contact surface (C). The characteristics of the strong-shock solutions are that at the contact surface with $\eta \cong 0.84$, $V = 1$, $\pi = 0$ and $\sigma \to \infty$; and that the mass efflux from the source can be expressed as

$$Q = 4\pi f \alpha \rho_0 r_0^2 \tag{7}$$

where $f \approx 0.4$ in the case of strong shocks. With V, σ and π known for $\eta \geq 0$, the physical quantities u, ρ and p can be derived in turn for various values of r and t from equation (5).

Details of the similarity solutions for a blast wave can be found in ref. 8 and numerical solution of the above adiabatic non-steady flow equations in ref. 9. For brevity, only radial variation of the mass density will be discussed here. Figure 1 shows the profile of the gas density (ρ) plotted for the case of strong forward (S_+) and reverse (S_-) shocks. As $\sigma \to \infty$ at the contact surface (C), ρ increases to a large value also. For weaker shocks, the snow-plow effect of piling up the material between S_+ and S_-, though less pronounced, should be present. The similarity solutions of the piston gas apply only to the inner coma region which is still dominated by collisional effect. For a Halley-type comet with a nuclear radius (r_n) of 3 km and a surface number density (n_0) of 3×10^{12} cm^{-3}, the nominal dimensions of such collision zone (r_c) is $\approx 1.4 \times 10^4$ km. As a result of the varying outgassing rate, r_c could be a factor of 3 larger or one-third smaller as the comet rotates around. Even with the large value of 4×10^4 km, r_c is still small compared with the radius of outer haloes ($\sim 3 \times 10^5$ km) (see ref. 1). The compression effect by the piston gas must then stop a long way inside. However, the sharp definition of the haloes may still be maintained at large distances if the temperature of the gas at the contact surface is kept sufficiently low by collisional cooling. The delineation of the haloes is also helped by the rare-fraction region just behind the reverse-shock S_-. Indeed, the density variations as shown in Fig. 1 may explain why in some cases a dark zone is observed[4] immediately behind a halo of enhanced brightness. Another effect is that the propagation speed of these haloes could be considerably lower than the value of 1 km s^{-1} generally quoted as the flow speed of the neutral gas in the inner coma close to the nucleus ($r \leq 10^{2-3}$ km) could be very small.

The dimension of these haloes as observed in the C_2-emission are limited by the scale length of these molecules ($\approx 10^5$ km at 1 AU). Furthermore, the daughter fragments of the parent molecules (such as H_2O and CO) in the compressed shells would pick up extra kinetic energy during photodissociation or ion-molecular reactions such that the corresponding thin-shell structures would be dispersed. Therefore, expanding haloes are not likely to be seen in the density distributions of H, C and O atoms, for example. We may then envisage the formation of a new spherical halo moving outwards every time the active hemisphere of the cometary nucleus faces the Sun. Partly due to the pile-up of material between the forward and reverse shocks, and partly due to the projection effect along the line-of-sight, a concentric halo will be observed progressing outwards. Note that the time variation of the outgassing rate could be anything but a step function as adopted here and the shock velocity α should have certain time variation. The dynamical process, however, is not likely to vary much as long as the active hemisphere of the comet can produce a strong driver gas. The real issue is, therefore, what is the energy source for such outbursts if they indeed are explosive.

In connection with the flare-ups of cometary activities sometimes observed, Donn and Urey[10] have suggested that explosive chemical reactions involving free radicals or unstable molecules on the nuclear surface may provide the answer. The free-radical reactions studied in the laboratory are observed to be highly exothermic, yielding a heat of reaction of approximately a few electron volts per molecule. This would mean that there are pockets of unstable chemical mixture on the surface of the active hemisphere with energetic explosions triggered by solar radiation. Another possibility suggested by Greenberg[11] in the case of interstellar grains—is simply that some of the solid grains could release free energy stored in the mantle of organic molecules and free radicals through breakup of the cometary grains. Although highly uncertain, it is, nevertheless, interesting to look into the energy budget of the coma expansion. Taking the radial speed to be 1 km s^{-1}, the kinetic energy is then only of the order of 0.1 eV per H_2O molecule and about $\dot{E} \approx 10^{28}$ eV s^{-1} in total if the gas production rate is $\approx 10^{29}$ H_2O molecules s^{-1}. An equivalent amount of chemical-energy source could be obtained if the sublimating ice contains 3% or so of the highly volatile material capable of exothermic chemical reaction. The energy budget of the halo expansion is therefore not a serious problem. In other words, even if the gas production rate were the same over one spin period, the contamination of one side with exothermic material would result in high-speed ejecta driving a shock pair periodically.

The hydrodynamics of the neutral coma in the inner region is not well understood. Various possible flow patterns including shock formation in steady-state conditions have been considered (see refs 12 and 13). The time-dependent expansion of the

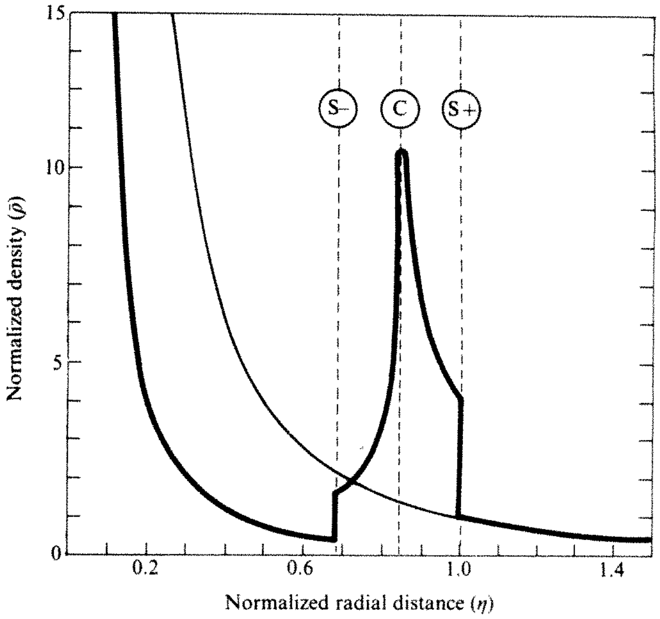


Fig. 1 The halo formation in the cometary comae through the compression effect of a piston shock. Radial variations of the density ($\bar{\rho}$) of the cometary neutral atmosphere normalized to the corresponding value at the position of the forward shock (S_+) at $\eta = 1$. Therefore $\bar{\rho} = \sigma(\eta)/\eta^2$ for $\eta = 1$ and $\bar{\rho} = 1/\eta^2$ for $\eta > 1$. The value of $\sigma(\eta)$ is obtained by taking a similarity solution to equations (1)–(6) with strong shocks (see ref. 8). The brightness enhancement in the haloes is caused by the density compression between the forward and reverse shocks. The dark zone sometimes observed immediately behind the haloes could be the manifestation of the rare-fraction region behind S_- as indicated. The density profile not modified by the shock pair is depicted by the thin curve.

cometary atmosphere should be even more complicated. Some idea can be derived from Bobrovnikoff's record of Comet Halley[1] where the expansion speed of the halo is estimated to be ≈ 0.25–$0.5\ \mathrm{km\,s^{-1}}$. Identifying this value as the propagation speed of the contact surface and using the relation that the flow speed u^* behind the reverse shock should be larger by a factor of 2.5[8], we obtain $u^* \approx 0.65$–$1.3\ \mathrm{km\,s^{-1}}$. This would mean that an additional energy source amounting to $\dot{E}^* \approx \frac{1}{2}Q(u^*)^2 \approx 10^{28}\ \mathrm{eV\,s^{-1}}$ (for $Q \approx 10^{29}$ H_2O molecules $\mathrm{s^{-1}}$) is needed to drive the piston gas. This requirement can be accommodated by introducing a small fraction of exothermic volatile material into the sublimating gas.

Whether any of these mechanisms may provide the energy required to drive a shock through the cometary atmosphere remains to be seen. But high-speed HCN jets ($u \geq 5\ \mathrm{km\,s^{-1}}$) have been detected in the coma of Comet Kohoutek (1973f) by Huebner *et al.*[14], who also favoured explosive chemical reactions as the origin of these jets. Thus the formation of shock pairs together with a compression region in between must be considered a viable explanation for the expanding haloes. For illustration and simplicity we have considered only the strong-shock case for the similarity solutions. Other types of compression processes with weaker shocks must also exist; the present interpretation, nevetheless, highlights the nature of non-steady expansion of the cometary atmosphere.

Received 28 July; accepted 14 November 1980.

1. Bobrovnikoff, N. T. *Univ. Calif. Publ. Lick Obs.* **17,** 2 (1931).
2. Whipple, F. L. *Astrophys. J.* **111,** 375 (1950).
3. Sekanina, Z. *Icarus* **37,** 420 (1979).
4. Rahe, J., Donn, B. & Wurm, K. *Atlas of Cometary Forms* (NASA SP-198, 1969).
5. Whipple, F. L. *Nature* **272,** 134 (1978).
6. Whipple, F. L. *IAU Circ.* No. 3459 (1980).
7. Courant, R. & Friedrichs, K. O. *Supersonic Flow and Shock Waves* (Interscience, New York, 1948).
8. Simon, M. & and Axford, W. I. *Planet. Space Sci.* **14,** 901 (1966).
9. Hundhausen, A. J. *J. geophys. Res.* **78,** 1528 (1973).
10. Donn, B. & Urey, H. C. *Astrophys. J.* **123,** 339 (1956).
11. Greenberg, J. M. *Moon & Planets* **20,** 15 (1979).
12. Shul'man, L. M. in *Physics of Comets* (ed. Konopleva, V. P.) 84 (transl. NASA TT F-599, 1969).
13. Mendix, D. A. & Ip, W.-H. *Astrophys. Space Sci.* **39,** 335 (1976).
14. Huebner, W. F., Snyder, L. E. & Buhl, D. *Icarus* **23,** 580 (1974).

Blocking highs over Asia and monsoon droughts over India

C. R. V. Raman

MONEX, New Delhi-3, India

Y. P. Rao

B3/33 Azad Apartments, New Delhi-3, India

Severe summer droughts of the Indian subcontinent have been found to accompany prolonged 'breaks' in the south-west monsoon. Data from the years of severe drought suggest that the associated breaks were due to upper tropospheric blocking ridges over East Asia. We present here a model for the evolution of two such blocking ridges: the East Asia (EABR) and the West Asia (WABR). Apparently an intense WABR—the initiator of the monsoon break—precedes the formation of the EABR. This cycle could be a way of predicting years of intense drought or monsoon activity.

For a long time, studies of monsoon breaks were confined to changes in circulation in the monsoon area. Ramaswamy[1–3] was the first to associate breaks with the low index circulation in middle latitude westerlies and the elongation of mid-tropospheric westerly troughs into India. However, eastward-moving westerly troughs penetrating into northern India enhance monsoon activity, and Ramanathan[4] had visualized that migrating middle latitude ridges influence monsoon onset. Examining the severe droughts of 1979 (MONEX), 1972, 1966 and 1965, Raman *et al.*[5] found that the breaks in these years were associated with stagnant upper tropospheric blocking ridges over the east Asia along ~100°E, extending from 35°N to 70°N. In such situations, westerly troughs do not move beyond 80°E and instead tend to extend southwards.

Cold lows form south-west of the EABR and the lower tropospheric monsoon trough over India gets displaced northwards into the Himalayas—the most significant characteristic of breaks. We suggest that an upper trophospheric WABR along ~50°E is a necessary precursor to the development of the EABR. Also that in drought years, a well-marked surface high appears over USSR centred at 55°N and 50°E. This surface high is linked by a ridge with the Azores High over the central Atlantic. The surface high over the USSR, overlain by a ridge of up to 100 mbar, is due to the WABR. In contrast, in good monsoon years, the surface high and the associated upper ridge over the USSR are notably absent. Instead, a trough from the Aleutian low extends into Siberia with several low centres. Such a configuration is akin to the normal pattern.

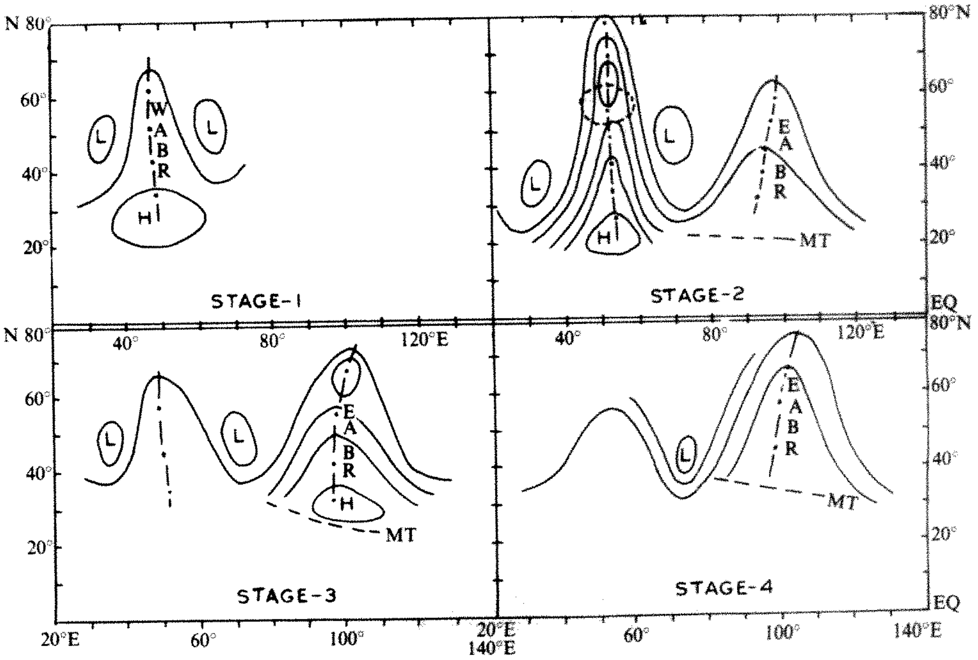


Fig. 1 Evolution of the WABR and the EABR. Solid lines denote 200-mbar contours; H, high; L, low; dashed-dotted line indicates north–south orientation of the 200-mbar blocking ridge. Stage 1, Appearance of WABR. Stage 2, Maximum phase of WABR with intense cold lows on either side. Surface high delineated by pecked isobar. EABR had commenced formation a few days earlier. Lower tropospheric Monsoon Trough (MT) in normal position. Stage 3, EABR had intensified along with weakening of WABR. Low upstream of the EABR shifts southwards. Stage 4, Monsoon trough is displaced significantly northwards leading to monsoon break.

Table 1 Evolution of WABR and EABR

		WABR (date of)			EABR (date of)		
Month and year	Duration of monsoon break/ deficient monsoon rains	Appearance (mbar)	Max intensity with value of closed contour at 300 mbar/ 200 mbar/ (gpm)*	Formation of surface high with departure from normal (mbar)	Appearance (mbar)	Max intensity with value of closed contour at 300 mbar/ 200 mbar (gpm)*	Formation of surface high with departure from normal (mbar)
July 1972	18–31 July	4 July (200)	14 July (1,240)	14 July (+14)	8 July (200)	18 July (1,200)	25 July (+14)
August 1972	1–9 August	29 July (200)	4 August (1,208)	8 August (+15)	1 August (200)	5 August (1,216)	8 August (+15)
August 1966	23–27 August	12 August (300)	20 August (944)	23 August (+12.5)	18 August (300)	23 August —	23 August (+23)
August 1965	4–15 August	25 July (300)	30 July (928)	4 August (+12.5)	31 July (300)	10 August (944)	10 August (+27)
July 1960	6–21 July	27 June (300)	1 July (944)	4 July (+12.5)	2 July (300)	5 July (952)	8 July (+20)

* gpm, Geopotential metres.

The evolution of the WABR and the EABR, during drought years from 1960, was obtained from Northern Hemispherical Synoptic charts[6]. The sequence of events preceding a monsoon break starts with the appearance of a predominant ridge at 200 mbar, first in the longitudinal belt 40–70°E, extending from the subtropical high pressure cell in the monsoon regime to almost 80°N. Stagnating at times for periods of up to one to two weeks, the WABR builds up as a warm high, reaching almost to the surface in a week after appearance. At its intense phase, even a closed high develops in the whole column, with the surface pressure departing from the long term normal (1931–60) by as much as +20 mbar. Such a deep meridionally-oriented blocking ridge configuration over west Asia elongates the westerly troughs both on the upstream (west European) and the downstream (central Asian) sides. About a week after the maximum phase, the WABR begins to weaken.

The EABR generally appears in the belt 90–120°E some four days after the WABR forms, and extends from 35 to 70°N. The EABR begins to build up in a similar way to the WABR, generally when the latter starts to shrink in meridional extent and weaken. During this phase of the build up of the EABR, the westerly trough over central Asia between the WABR and the EABR further elongates to more southern latitudes. (The temperatures in this trough are ~5–8 °C below normal in the mid-troposphere but warmer than the surrounding at 200 mbar and 100 mbar.) The superposition of this elongated cold upper trough/low at southern latitudes is manifested as the displace-

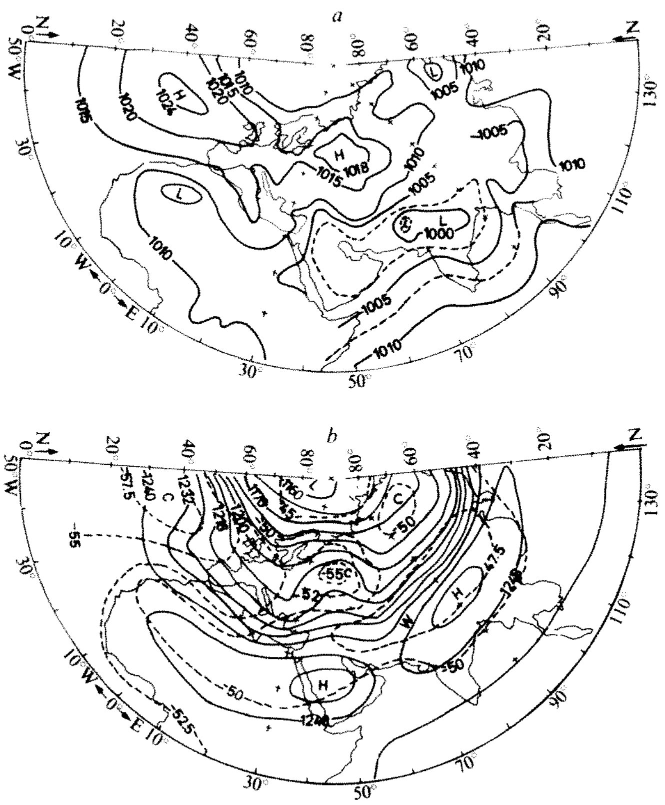


Fig. 2 Mean conditions in August 1972, a year of severe monsoon drought. *a*, Surface isobars. Note marked high centre at 55°N 50°E and the associated ridge linking with the Azores High over the central Atlantic. The ridge extends to 90°E across Eurasia. *b*, 200 mbar. Solid lines denote contours and dashed lines represent temperature (°C). C, cold; W, warm. Note markedly cold upper tropospheric ridge along 50°E, above the surface high over USSR.

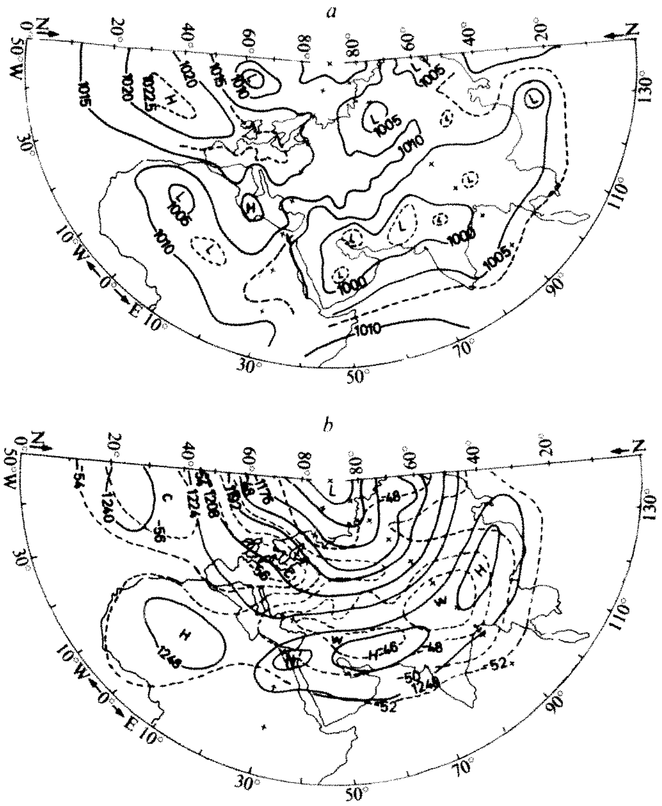


Fig. 3 Mean conditions during August 1975, a year of good monsoon rains. *a*, Surface isobars. Note the absence of a high centre over the USSR. Instead, a trough from the Aleutian low extends into Siberia with several low centres. Note that the monsoon low over India and neighbourhood is more intense than in Fig. 2*a* (1972). *b*, 200 mbar. Solid lines denote contours and dashed lines represent temperature (°C). Note the circumpolar flow and absence of ridges along 50°E.

ment of the lower tropospheric monsoon trough (MT) to the north into the Himalayas—symptomatic of the monsoon break. The build-up of an intense WABR seems essential for creating the EABR, which brings about the monsoon break. This cycle of the WABR building up well before the break could be useful in predicting monsoon breaks and consequent droughts. The sequence of evolution is presented in Fig. 1. A few examples of the cycle are given in Table 1.

Certain broad features of the WABR and the EABR are striking even in the monthly mean charts of 1972, a severe drought year. Figures 2 and 3 present the surface and upper air circulation features during August 1972 and 1975 respectively. In 1972, the WABR had started to build-up by the beginning of June and continued during the first half of July 1972. The subsequent build up of the EABR coincided with the pronounced break over India during the second half of July. The WABR/EABR configuration continued in August 1972, when the monsoon rains failed. The WABR prominent in Fig. 2*b* at 200 mbar, is associated with cold temperatures along ~50°E. The surface high vertically below this 200 mbar ridge is quite prominent (Fig. 2*a*). The EABR is less marked in the monthly mean chart (Fig. 2*b*), although it is seen markedly in decadal mean charts covering break occurrences in this and other years. In the good monsoon year of 1975, the circulation at 200 mbar (Fig. 3*b*) is more circumpolar with a weak trough in the location of the 1972 WABR (50°E). The contrasts between the good and bad monsoon years described above in the surface pressure distribution and the upper air features from the eastern Atlantic to east Asia are shown by other years during the past two decades.

Negative temperature anomalies are noticed at 200 mbar and above in WABR and EABR, with positive anomalies in the troposphere. The tropospheric warmth seems to be associated with the subsidence characteristic of high pressure. The stratospheric cooling must be due to the ascent, and maintains the high. Hoskins and Simmons have revived interest in downstream development as an important atmospheric scale phenomenon. It might be useful to apply this study to monsoon variability over India.

Received 4 August; accepted 14 November 1980.

1. Ramaswamy, C. *Geophysica* **6,** 455–477 (1958).
2. Ramaswamy, C. *Tellus* **14,** 337–349 (1962).
3. Ramaswamy, C. *Bull. Ind. natn. Sci. Acad.* **54,** 109–132 (1975–76).
4. Ramanathan, A. S. *J. mar. Biol. Ass. Ind.* **14,** 843–861 (1972).
5. Raman, C. R. V., Rao, Y. P. & Alvi, S. M. A. *Curr. Sci. Ind.* **49,** 123–129 (1980).
6. *Synoptic Bull. Northern Hemisphere* Parts I–III (USSR, 1960–1978).
7. Simmons, A. J. & Hoskins, B. J. *J. atmos. Sci.* **36,** 1239–1254 (1979).

Possible role for metals in stratospheric chlorine chemistry

Edmond Murad & William Swider

Air Force Geophysics Laboratory, Hanscom Air Force Base,
Massachusetts 01731

Sidney W. Benson

Department of Chemistry, University of Southern California,
Los Angeles, California 90007

Meteoroids ablate largely in the Earth's ionosphere (at altitudes between 80 and 120 km) and the products are seen mostly as atomic ions[1–3]. Atomic neutral atoms have also been observed and studied optically, but the data are not extensive for elements other than sodium[4,5]. The processes responsible for maintaining the equilibrium of metallic atoms and ions in the ionosphere are understood generally[6–9]. Below 80 km, the metal ions are largely neutralized. The chemistry which occurs during the subsequent transport of the metals to the ground is uncertain; aerosols collected in the lower stratosphere have been reported to contain chlorides and sulphates of metals (some of which occur in meteors)[10–12] while aerosols collected in the troposphere have been found to contain many metals among which are meteor metals[13–16]. It has been suggested that ions of the type $Na^+ \cdot (H_2O)_n$ may be present in the upper stratosphere[17], their presence being plausible in view of a meteor sodium model which indicated that large amounts of NaOH (almost 10^6 cm^{-3}) would be present at 40 km (ref. 18). This latter conclusion was questioned[19] in view of thermochemical data which indicated that NaOH, once formed, should be removed quickly by reaction with Cl, ClO and HCl. We consider here reactions of meteor, and other, metals with Cl, ClO and HCl to assess the impact of halocarbons on ozone. Release of 1,000 tonne of sodium (or similar metal) may significantly reduce the harmful effects of halocarbons on ozone for several years.

Although sodium comprises only about 2% by weight of the total metal concentration in meteors[20], its reaction with Cl, ClO and HCl may represent a class of reactions which could be important in the chemical maintenance of these species in the stratosphere (the condensation of Ga, Ge and Sb in nebulae has been suggested to proceed through the respective metal hydroxide and metal chloride intermediates[21]). The other major metallic constituents of meteors are (by weight): Al(1.7%), Ca(1%), Fe(11.5%), Mg(12.5%), Ni(1.5%) and Si(20%). Here Si is included in the term 'metallics'. Other, less abundant metals (such as Ti, Cr, Mn) have been observed in micrometeorites collected in the stratosphere[22].

We now consider chemical reactions involving these metals or their intermediates with Cl, ClO and HCl, even though a quantitative evaluation of these reactions is not possible because of the lack of kinetic and thermochemical data. Cl, ClO and HCl have significant concentrations at altitudes between 15 and 50 km (refs 23, 24). Sodium will probably be in the form of NaOH (ref. 18) and/or $NaHCO_3$ by the time it has descended to 50 km (refs 19, 25). Similar compounds will be formed in the case of the other metals. The basic mechanism for the formation of the metal hydroxides (MOH) and the bicarbonates ($MHCO_3$) seems to be:

$$MO + H_2O \rightarrow MOH + OH \quad (M = \text{Na and, possibly, Al}) \qquad (1)$$

$$MO + CH_4 \rightarrow MOH + CH_3 \quad (M = \text{Na, Al}) \qquad (2)$$

$$MO + H_2 \rightarrow MOH + H \quad (M = \text{Na, Al}) \qquad (3)$$

$$MOH + CO_2 + X \rightarrow MHCO_3 + X \quad (\text{for all MOH}) \qquad (4)$$

where M is any meteor metal and X is any third-body. Metal chlorides, MCl, may then be formed by the reactions:

$$MOH + Cl \rightarrow MCl + OH \quad (M = \text{Na}) \qquad (5)$$

$$MOH + ClO \rightarrow MCl + HO_2 \quad (\text{all M}) \qquad (6)$$

$$MOH + HCl \rightarrow MCl + H_2O \quad (\text{all M}) \qquad (7)$$

$$MHCO_3 + HCl \rightarrow MCl + (H_2CO_3 \text{ or } H_2O + CO_2) \quad (\text{all M}) \qquad (8)$$

Reactions (7) and (8) are expected to proceed with large rate coefficients (~10^{-9} cm^3 molecule^{-1} s^{-1}) and can deplete HCl significantly relative to the reaction

$$HCl + OH \rightarrow H_2O + Cl \qquad (9)$$

which is an important step in the chain reaction leading to the depletion of ozone[26]. Reaction (9) has a rate coefficient[27] of 6.6×10^{-13} cm^3 molecule^{-1} s^{-1} so that even if $k_7 + k_8$ [rate coefficients for reactions (7) and (8)] were a factor of 1,000 less than estimated (that is ~10^{-12} cm^3 molecule^{-1} s^{-1}), they would still be important competition for reaction (9) for [OH]~[MOH]+[$MHCO_3$]; the data of refs 18 and 23 indicate this to be true at ~20 km. Of the meteor metals, sodium and aluminium

0028–0836/81/040273—02$01.00

can readily form chlorides through reactions (7) and (8). The possibility of forming hydroxides of other metals (Fe, Mg and Ca) by three-body reactions or by reactions of the monoxides with minor constituents (such as HO_2) should not be discounted.

Silicon is likely to enter the stratosphere as SiO_2 as the large dissociation energy of the monoxide, $D^\circ_0(\text{Si–O}) = 190 \pm 2\ \text{kcal mol}^{-1}$ (ref. 28), precludes its reduction by O atoms. SiO_2 might also form weakly bonded compounds with Cl, ClO and HCl; photochemical studies indicate that solid SiO_2 catalyses the rate of photochemical decomposition of halocarbons[29]. Reactions leading to the formation of MOCl may also be possible, perhaps through

$$\text{MO} + \text{ClO} \rightarrow \text{MOCl} + \text{O} \qquad (10)$$

Thermochemical data are available only for AlOCl, and these indicate that the dissociation energy of AlOCl, $D^\circ_0(\text{AlO–Cl})$ is greater than $D^\circ_0(\text{Cl–O})$ or $D^\circ_0(\text{H–Cl})$ (ref. 30). Indeed, analysis of the high-temperature reactions among the constituents of solid-fuel rocket engines predicts large amounts of AlOH and AlOCl at the expense of Al_2O_3 and HCl (ref. 31) (note that the engine of the space shuttle releases considerable amounts of Al_2O_3, Fe and $FeCl_2$ (ref. 32)).

A quantitative assessment of the importance of reactions (1)–(8) will have to await the availability of their rate coefficients. However, the strongly ionic character of many of these metals suggests that reactions of the above type will be fast. For example, Ca and ClO_2 react to yield both CaO and CaCl and the total reaction cross-section has been estimated to be $10^{-14}\ \text{cm}^2$ at thermal energies[33], which corresponds to a rate coefficient of $10^{-9}\ \text{cm}^3\ \text{molecule}^{-1}\ \text{s}^{-1}$. Taking $k_7 + k_8$ to be $10^{-11}\ \text{cm}^3\ \text{molecule}^{-1}\ \text{s}^{-1}$ (which is a conservative estimate) and using $[\text{MHCO}_3] + [\text{MOH}] \sim 2 \times 10^6\ \text{cm}^{-3}$ for M = Al + Na at 40 km [this is equivalent to twice the calculated total sodium compound concentration at 40 km (ref. 18)], then the time constant for reactions 7 + 8 is 5×10^4 s. On the other hand, reaction (9), the dominant process for the removal of HCl not involving metals, has a time constant of $\sim 2 \times 10^6$ s because $[\text{OH}] \sim 10^6\ \text{cm}^{-3}$ and $k_9 = 6.6 \times 10^{-13}\ \text{cm}^3\ \text{molecule}^{-1}\ \text{s}^{-1}$ (ref. 27). This shows that reactions such as (7) and (8) may be important.

The metal chlorides, particularly NaCl, are very stable in the stratosphere, and would be expected to act as nucleating agents, eventually being incorporated in aerosols. (Rocket exhaust which contains metals such as aluminium has been found to enhance ice nucleation[34].) Observations indicate the presence of Cl (refs 10–12, 16, 35), and Na in some aerosols. At high temperatures both condensed and gaseous Na_2SO_4, a possible aerosol constituent, reacts[36,37] with HCl to yield NaCl. NaCl will, of course, polymerize readily. The condensation of NaCl will probably be in the form of heterogeneous nucleation of monomer, dimer or trimer NaCl with other constituents such as water molecules or sulphur dioxide. A cluster containing 100 molecules, only 1–3 of which are NaCl, will have a concentration of $\sim 10^6\ \text{cm}^{-3}$ and its sodium would be undetectable. On the other hand, a cluster of NaCl, $(\text{NaCl})_n$ with $n = 100$, would have a concentration of $\sim 10^4\ \text{cm}^{-3}$ and would again be undetectable in the atmosphere.

How much metal is deposited by meteors and how does this compare with the total amount of halocarbons deposited in the atmosphere? Parkin and Tilles[38] estimated the total meteor influx to the Earth to be from $3.6 \times 10^{10}\ \text{g yr}^{-1}$ to $3.6 \times 10^{11}\ \text{g yr}^{-1}$. Gadsden[7] estimated the total meteor influx to be $9 \times 10^{10}\ \text{g yr}^{-1}$ (9×10^4 tonnes per yr), corresponding to a total metal atom influx of 2×10^6 metal atoms $\text{cm}^{-2}\ \text{s}^{-1}$. Hughes[39] quoted values in the range 3×10^8–$5 \times 10^{13}\ \text{g yr}^{-1}$ with a recommended estimate of $5 \times 10^8\ \text{g yr}^{-1}$. Later, Hughes[40] reviewed the data and estimated the total amount of meteors deposited into the Earth to be $1.6 \times 10^{10}\ \text{g yr}^{-1}$. Hunten *et al.*[41] analysed this problem from the point of view of smoke and dust particles deposited in the Earth's mesosphere and stratosphere by meteoroids and preferred the value given by Hughes[40], although the possibility of a higher value was noted, and suggested[41] a meteoric origin for the condensation nuclei of sulphate aerosols. The total weight of halocarbons deposited in the atmosphere has been estimated to be $\sim 7 \times 10^5$ tonnes yr^{-1}. If Gadsden's estimate[7] is correct, then meteor metals may be important in the chemistry and fate of the halocarbons in the stratosphere.

High-flying aircraft provide two additional sources of metals in the stratosphere through the combustion of jet fuels and the erosion of engine parts (the latter being orders of magnitude greater than the former[42]). The principal metal contaminants in the fuels are iron, copper, calcium, vanadium and sodium, and these may be released in amounts varying between 100 and 1,000 tonnes yr^{-1} depending on the type of fuel and on the frequency of flights[43]. Engine erosion elements include predominantly aluminium, manganese and chromium.

An intentional or an unintentional release (such as, jet fuels and engine erosion) in the stratosphere of $\sim 10^3$ tonnes of Na and, possibly, Al might reduce [Cl] and [ClO] by a factor of two. It would require several years for the diffusion cycle to re-establish those concentrations if the release occurred above 30 km. Finally, although the above discussion is speculative, it may help explain recent controversial results obtained by Anderson *et al.*[44], who were the first to measure [Cl], [ClO] and $[O_3]$ simultaneously. Their results indicate some difficulties in current chemical models used to predict the effect of these species on $[O_3]$.

We thank Dr R. S. Narcisi for helpful comments.

Received 16 April; accepted 15 November 1980.

1. Narcisi, R. S. *Space Res.* **8,** 360–369 (1968).
2. Zbinden, P. A., Hidalgo, M. A., Eberhardt, P. & Geiss, J. *Planet. Space Sci.* **23,** 1621–1642 (1975).
3. Goldberg, R. A. & Aikin, A. C. *Science* **180,** 294–296 (1973).
4. Hunten, D. M. *Space Sci. Rev.* **6,** 493–573 (1967).
5. Megie, G., Bos, F., Blamont, J. E. & Chanin, M. L. *Planet. Space Sci.* **26,** 27–35 (1978).
6. Swider, W. *Planet. Space Sci.* **17,** 1233–1246 (1969).
7. Gadsden, M. *Annls Geophys.* **26,** 141–150 (1970).
8. Brown, T. L. *Chem. Rev.* **73,** 645–667 (1973).
9. Murad, E., *J. geophys. Res.* **83,** 5525–5530 (1978).
10. Shedlovsky, J. P. & Paisley, S. *Tellus* **18,** 499–503 (1966).
11. Delaney, A. C., Shedlovsky, J. P. & Pollock, W. H. *J. geophys. Res.* **79,** 5646–5650 (1974).
12. Lazrus, A. L. & Gandrud, B. *J. geophys. Res.* **79,** 3424–3431 (1971).
13. Cadle, R. D. & Grams, G. W. *Rev. Geophys. Space Phys.* **13,** 475–501 (1975).
14. Charlson, R. J., Covert, D. S., Larson, T. V. & Waggoner, A. P. *Atmos. Envir.* **12,** 39–53 (1978).
15. Penkett, S. A., Atkins, D. H. & Unsworth, M. H. *Tellus* **31,** 295–307 (1979).
16. Parungo, F., Ackerman, E., Caldwell, W. & Weickmann, H. K. *Tellus* **13,** 521–529 (1979).
17. Ferguson, E. E. *Geophys. Res. Lett.* **5,** 1035–1038 (1978).
18. Liu, S. C. & Reid, G. C. *Geophys. Res. Lett.* **6,** 283–286 (1979).
19. Murad, E. & Swider, W. *Geophys. Res. Lett.* **6,** 929–932 (1979).
20. Dube, A. *et al.* in *Mineral Science Investigations, 1974–75* (ed. Mason, B.) (Smithsonian Contributions to the Earth Sciences No. 19, Washington DC, 1975).
21. Wai, C. M. & Wasson, J. T. *Nature* **282,** 790–793 (1979).
22. Ganapathy, R. & Brownlee, D. E. *Science* **206,** 1075–1077 (1979).
23. Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. *Phil. Trans. R. Soc.* **290,** 187–234 (1978).
24. Abbas, M. M. *et al. J. geophys. Res.* **84,** 2681–2690 (1979).
25. Keneshea, T. J., Murad, E., Swider, W. & Zimmerman, S. P. *EOS* **60,** 898 (1979).
26. Rowland, F. S. & Molina, M. J. *Rev. Geophys. Space Phys.* **13,** 1–35 (1975).
27. Baulch, D. L. *et al. J. phys. Chem. Ref. Data* **9,** 295–472 (1980).
28. Stull, D. R. & Prophet, H. (eds) *JANAF Thermochemical Tables* 2nd edn NSRDS-NBS-37 (US Government Printing Office, Washington DC, 1971).
29. Ausloos, P., Rebbert, R. E. & Glasgow, L. *J. Res. natn. Bur. Stds*, **82,** 1–8 (1977).
30. Chase, Jr., M. W., Curnutt, J. L., McDonald, R. A. & Syverud, A. N. *J. phys. Chem. Ref. Data* **7,** 311–480 (1978).
31. Park, C. *Atmos. Envir.* **10,** 693–702 (1976).
32. Isaacs, R. G., Burke, H.-H. & Sze, N. D. *Environmental Impact Studies for Several Space Transportation Vehicles* (Environmental Research and Technology, No. P-3139, Concord, Massachusetts, 1977).
33. Engelke, F., Sander, R. K. & Zare, R. N. *J. chem. Phys.* **65,** 1146–1155 (1976).
34. Parungo, F. P. & Allee, P. A. *J. appl. Met.* **17,** 1856–1863 (1978).
35. Lazrus, A. L., Gandrud, B. & Cadle, R. D. *J. geophys. Res.* **76,** 8083–8088 (1971).
36. Stearns, C. A., Kohl, F. J., Fryburg, G. C. & Miller, R. A. in *High Temperature Metal Halide Chemistry* (eds Hildenbrand, D. L. & Cubicciotti, D. D., 555–573 (Electrochemical Society, Princeton, 1977).
37. Hancock, P. in *High Temperature Metal Halide Chemistry* (eds Hildenbrand, D. L. & Cubicciotti, D. D., 645–669 (Electrochemical Society, Princeton, 1977).
38. Parkin, D. W. & Tilles, D. *Science* **159,** 936–946 (1968).
39. Hughes, D. W. *Planet. Space Sci.* **20,** 1949–1959 (1972).
40. Hughes, D. W. *Space Res.* **15,** 531–539 (1975).
41. Hunten, D. M., Turco, R. P. & Toon, O. B. *J. atmos. Sci.* **37,** 1342–1357 (1980).
42. *CIAP Monogr. 2, Propulsion Effluents in the Stratosphere,* DOT-TST-75-52 (NTIS, Springfield, 1975).
43. *CIAP Monogr. 3, The Stratosphere Perturbed by Propulsion Effluents,* DOT-TST-75-53 (NTIS, Springfield, 1975).
44. Anderson, J. G., Grassl, H. J., Shetter, R. E. & Margitan, J. *J. geophys. Res.* **85,** 2869–2887 (1980).

Humic soil and coal structure study with magic-angle spinning ^{13}C CP-NMR

P. F. Barron

CSIRO Fuel Geoscience Unit, North Ryde, New South Wales 2113, Australia

M. A. Wilson

Fuel and Atmospheric Chemistry Section, CSIRO Division of Process Technology, North Ryde, New South Wales 2113, Australia

In our recent report on using the ^{13}C cross-polarization (CP) NMR technique to characterize the organic matter in whole soils[1], we discussed the limitations of the technique in terms of resolution and the improvements possible using 'magic-angle' sample spinning (MAS). We now report experiments on a soil and a brown coal in which these techniques have been combined. The spectra provide information on the chemical structure of the materials and throw new light on the nature of the various types of carbon in coals and humic materials.

Although carbonaceous substances such as coals and humic materials are among the most widely distributed and important natural products on the Earth, there have been few techniques available by which their intimate chemical structure could be obtained. However, NMR techniques have allowed semi-quantitative estimations of aromatic carbon in coals[1-3] and whole soils[1] to be made. Recently, high power cross-polarization techniques coupled with magic-angle spinning[4] have improved quantification, through increased resolution, enabling better estimates of the aromaticity of coals[5] and humic substances[6]. High field techniques have also been used to obtain structural information on coals[7] and soils[8], but research here has had to be confined to soluble material.

We have combined the ^{13}C CP-MAS technique with the use of high fields to obtain spectra at 75 MHz for a New Zealand silt loam and an Australian brown coal xylite fraction. The spectra (Fig. 1) consist of many broad resonances, each of which gives information on the structural environments of the carbon atoms present.

A major problem with obtaining ^{13}C CP-MAS spectra is the need to spin at speeds sufficient to remove both the chemical shift anisotropy (CSA) of highly anisotropic carbons (that is, aromatic and carboxyl carbons) and the spinning sidebands from the chemical shift range. As the CSA is field dependent, the required spinning speed increases with increasing field strength and speeds of the order of the CSA are usually considered necessary. At 75 MHz, experiments would have to be performed at speeds in excess of 7 kHz, which are not technologically feasible at present. However, Stejskal *et al.*[9] have shown that an anisotropy of ~3 kHz is effectively removed at spinning speeds of ~1 kHz but at the expense of an increase in spinning sideband intensities. The spinning speeds we have used at 75 MHz (3.2–3.6 kHz) should be sufficient to remove large anisotropies effectively and hence reveal the expected resolution advantage of high fields. The increased problem of spinning sidebands is difficult to solve but, because their position depends on the spinning rate, their contribution can be reduced by multiplication of spectra obtained at significantly different spinning speeds.

Figure 1*a* shows the ^{13}C CP-MAS spectrum of Maungatua silt loam. We have previously obtained a 22.6 MHz ^{13}C CP spectrum[1] which indicated that the silt loam was largely aliphatic and highly oxygenated. The high field CP-MAS (3.2 kHz) spectrum confirms these conclusions and provides greater detail. However, as a consequence of the low carbon content of the soil (22% C) the spectrum is complicated by resonances from the rotor (perdeuterated polymethylmethacrylate[9]).

^{13}C resonances in the spectrum can be assigned if it is assumed that the isotropic chemical shifts (δ) in such solids are not significantly (±3 p.p.m.) different from those found in solution[10,11]. The spectrum can be divided into four regions: δ 10–50 p.p.m. (aliphatic carbon), δ 50–110 p.p.m. (oxygenated alkyl carbon), δ 100–160 p.p.m. (aromatic and alkenic carbon) and δ 160–220 p.p.m. (carbonyl carbon).

The strong group of resonances at ~δ 176 p.p.m. arises from carboxyl carbon in the soil and/or the rotor. Clearly carboxyl carbon is only a minor contributor to this soil. The O-alkyl region exhibits signal over the range δ 50–110 p.p.m. with the resolved resonance at δ 74 p.p.m. indicating oxygen-bonded carbons in ethers other than methoxy or ethoxy, or possibly ring carbons in polysaccharides. That the O-alkyl region extends to greater than 100 p.p.m. would be consistent with the presence of polysaccharides[12]. Unfortunately, the methoxy and methyl ester region is complicated by strong rotor resonances (δ 55–51 p.p.m.) and therefore we cannot comment on the presence of these groups.

The alkyl region contains clearly defined resonances at δ 29–32 and 15 p.p.m. The range of signals δ 29–32 p.p.m. presumably arises from methylene carbon β, γ, δ and ε from terminal methyl groups. The signal at δ 15 p.p.m. arises from terminal methyl carbon of alkyl chains. A small but probable contribution from rotor signals to this resonance, means the only

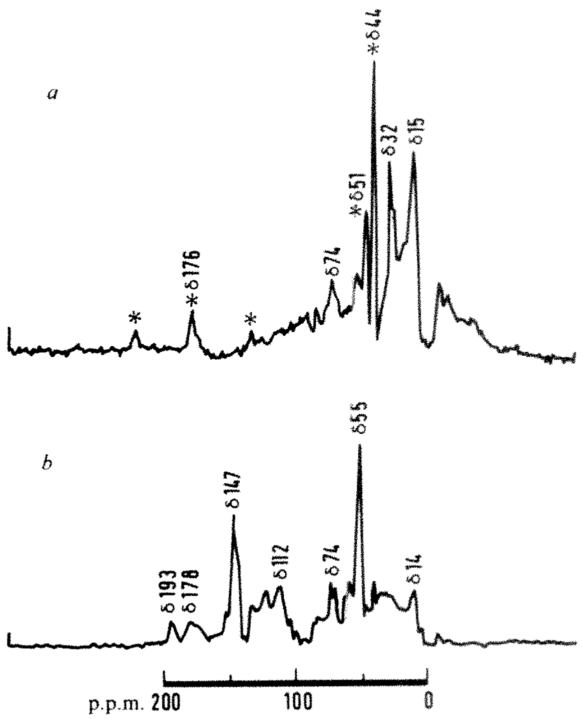


Fig. 1 75 MHz ^{13}C CP–MAS spectra obtained using H_1 fields of 19 and 76 G for ^1H and ^{13}C respectively, with a single contact of 1 ms, on a Bruker CXP-300 spectrometer and Z32 PE/MAS probe. The rotor was machined from perdeuterated polymethylmethacrylate. *a*, Maungatua silt loam. Obtained without block multiplication and using a spinning speed of 3.2 kHz. * Indicates rotor signals and sidebands. *b*, Yallourn brown coal xylite. Obtained by block multiplication of spectra using spinning speeds of 3.2 and 3.6 kHz to reduce spinning side bands. The resulting free induction decay was gaussian resolution enhanced. These materials have been described in detail elsewhere[1]. No special procedures were needed for sample preparation other than crushing. Assignments were made by comparison with known compounds[24,25]. Chemical shifts are relative to external tetramethylsilane and were initially determined relative to external adamantane and are considered accurate to ±3 p.p.m.

0028-0836/81/040275—02$01.00

comment we can make on alkyl chain length is that polymethylene is an important constituent of this soil. Our results clearly show, however, the potential of MAS in soil research. We are investigating the potential of other rotor materials for obtaining spectra of whole soils.

Numerous attempts have been made to devise structural formulae which represent the alkali soluble fractions of soil organic matter. These models[13–15] differ considerably in structure but all involve the idea that soil humic substances are highly aromatic materials.

It has recently been suggested that soil humic substances may not be as aromatic as previously thought. Anderson and Russell[16] suggested that polymaleic-acid-like polymers are important contributors to humic substances. Grant[17] has shown that polymethylene may constitute as much as 30% by weight of humic structures. Signals from aromatic carbons contribute only moderately to the ^{13}C NMR spectra of some soil humic materials in solution[18–20]. Although it may be argued[18] that nuclear Overhauser and relaxation effects prevent quantitative determination from ^{13}C NMR of the extract aromaticity, gated ^{1}H decoupling and significant pulse delays have little effect[20] on relative signal intensities of soil humics recorded so far. Our work on whole soils[1] has shown that the soil organic matter present can vary from highly aliphatic to at least moderately aromatic. The soil investigated in the present study is clearly highly aliphatic. Hence the structure models proposed by Fuchs[13], Felbeck[14], Dragunov[13] and Schnitzer[15] are not applicable to this soil, though they may adequately represent the organic matter in extracts from other soils.

Figure 1*b* is the spectrum obtained from a Yallourn brown coal xylite. Previous examination of this material[1] by the ^{13}C CP technique (without MAS) at the relatively low frequency of 22.6 MHz showed that the xylite contained considerable amounts of aromatic, O-alkyl and alkyl carbon. The presence of highly anisotropic aromatic carbon would be expected to lead to significant sideband problems. Hence, Fig. 1*b* is the result of the multiplication of spectra obtained at two spinning speeds to reduce this problem. The 75 MHz ^{13}C CP-MAS spectrum exhibits considerable detail and confirms the previous conclusions.

The broad strong resonances present in the spectrum are clearly not from spinning side bands. The resonance at δ147 arises from aromatic carbons bound to oxygen whereas that at δ112 is indicative of carbon *ortho* to C_{Ar}–O or of alkenic carbons. The major resonance at δ123 p.p.m. is due to C_{Ar}–H carbons.

The region δ50–100 p.p.m. can be assigned to the O-alkyl carbon of alcohols, ethers and esters. Resonance from δ70–90 p.p.m. can in general be assigned to carbon α to oxygen where the alkyl group contains more than two carbons. The spectrum exhibits signals over this whole range. Primary alcohols, ethoxy ethers and ethyl esters resonate over the range δ60–65 p.p.m. and small amounts of these materials may be present. The strong resonance at δ55 p.p.m. can be assigned to methoxy carbons. The presence of large amounts of methoxy groups has been confirmed by pyrolysis–gas chromatography and mass spectroscopy.

The alkyl region extending from δ10 to δ50 p.p.m. is not dominated by any alkyl types. There is no large polymethylene peak at δ29 p.p.m. Zilm *et al.*[5] obtained the 15.1 MHz ^{13}C CP–MAS spectrum of an American lignite and identified a major peak at δ29 p.p.m. which is readily assigned to polymethylene $(CH_2)n$. There is only a very minor peak at δ29 p.p.m. in the xylite spectrum. It follows that the alkyl chains in Yallourn brown coal xylite are short and branched whereas the alkyl chains in the American coal are much longer. Hence our results indicate the variability of alkyl structures in brown coal components.

Assigning resonances to carboxyl and carboxylic carbon is complicated by the uncertainty of the contribution of spinning side bands to the spectrum but brown coals are known to contain carboxylic carbon. Although the peak at 165–180 p.p.m. could be assigned to carboxylic carbon it may also arise from a residual spinning side band signal. The peak at δ193 p.p.m. possibly represents ketone carbon, although residual spinning side bands may contribute to this region. Ketone and carboxylic carbon were, however, also identified in pyrograms.

It is generally believed[21,22] that brown coals contain considerable amounts of carboxylic acid, ether and phenolic functional groups bound to aromatic structures of one or two rings. Our results show the important contribution of methoxy ethers in brown coal xylite. The very strong resonance at δ55 p.p.m. highlights the importance of methoxy ethers in this brown coal. The unequivocal assignments made here (to those signals which cannot arise from spinning side bands) are also those found in lignin[23]. This strongly suggests that Yallourn xylite is a modified lignin.

We are not aware of any reported NMR spectra of whole soils or solid coal which permit the detailed structural assignments achieved here. The present results show the potential and problems of high field high power NMR techniques in soil and coal chemistry.

We thank Drs D. Muller and H. Forster, and Bruker Analytische Messtechnik GMBH, Karlsruhe, FRG for obtaining the CP-MAS spectra. P.F.B. was supported under the National Energy Research, Development and Demonstration program administered by the Commonwealth Department of National Development.

Received 7 July; accepted 3 November 1980.

1. Barron, P. F., Wilson, M. A., Stephens, J. F., Cornell, B. A. & Tate, K. R. *Nature* **286,** 585–587 (1980).
2. Retcofsky, H. L. & VanderHart, D. L. *Preprints 1976 Coal Chemistry Workshop*, 202–217 (Stanford Research Institute, Menlo Park, 1976).
3. Vanderhart, D. L. & Retcofsky, H. L. *Fuel* **55,** 202–204 (1976).
4. Schaefer, J. & Stejskal, E. O. *J. Am. chem. Soc.* **98,** 1031–1032 (1976).
5. Zilm, K. W., Pugmire, R. J., Grant, D. M., Wood, R. E. & Wiser, W. H. *Fuel* **58,** 11–16 (1979).
6. Hatcher, P. G., VanderHart, D. L. & Earl, W. L. *Org. Geochem.* **2,** 87–92 (1980).
7. Bartle, K. D. *Rev. Pure appl. Chem.* **22,** 79–113 (1972).
8. Wilson, M. A., Jones, A. & Williamson, B. F. *Nature* **276,** 487–489 (1978).
9. Stejskal, E. O., Schaefer, J. & McKay, R. A. *J. Magn. Reson.* **25,** 569–573 (1977).
10. Lippmaa, E. T., Alla, M. A., Pehk, T. J. & Engelhardt, G. *J. Am. chem. Soc.* **100,** 1929–1931 (1978).
11. Earl, W. L. & VanderHart, D. L. *Macromolecules* **12,** 762–767 (1978).
12. Voelter, W. & Breitmaier, E. *Org. magn. Reson.* **5,** 311–319 (1973).
13. Kononova, M. M. *Soil Organic Matter* 2nd edn, 73–75 (Pergamon, New York, 1976).
14. Felbeck, G. T. Jr *Adv. Agron.* **17,** 327–333 (1965).
15. Schnitzer, M. & Khan, S. Y. *Soil Organic Matter,* 45–47 (Elsevier, New York, 1978).
16. Anderson, H. A. & Russell, J. D. *Nature* **260,** 597 (1976).
17. Grant, D. *Nature* **270,** 709–710 (1977).
18. Wilson, M. A. & Goh, K. M. *J. Soil Sci.* **28,** 645–652 (1977).
19. Ogner, G. *Soil Biol. Biochem.* **11,** 105–108 (1979).
20. Newman, R. H., Tate, K. R., Barron, P. F. & Wilson, M. A. *J. Soil Sci.* (in the press).
21. Wender, I. *Catal. Rev. Sci. Engng* **14,** 97–129 (1976).
22. Given, P. H. *Fuel* **39,** 147–153 (1960).
23. Mikinis, F. P., Maciel, G. E. & Bartuska, V. J. *Org. Geochem.* **1,** 169–176 (1979).
24. Levy, G. C. & Nelson, G. L. *Carbon-13 Nuclear Magnetic Resonances for Organic Chemists* (Wiley-Interscience, New York, 1976).
25. Stothers, J. B. *Carbon-13 N. M.R. Spectroscopy* (Academic, New York, 1976).

Was the Laramide orogeny related to subduction of an oceanic plateau?

Richard F. Livaccari, Kevin Burke & A. M. C. Şengör

Department of Geological Sciences, State University of New York at Albany, Albany, New York 12222

Numerous models have been presented to explain the late Cretaceous/early Cenozoic Laramide orogeny, which affected the foreland region of the western cordillera within the US[1]. The most attractive models invoke low-angle subduction[2–4], which can develop for various reasons[5,6]. Evidence from South America indicates that one such reason may be the subduction of buoyant ocean floor. We here extend the low-angle subduction models by attributing the shallowing of the subduction angle during the Laramide orogeny to subduction of a large oceanic plateau of anomalously thick and buoyant oceanic crust. Pacific plate history indicates that a twin to the Hess Rise may represent the oceanic plateau that triggered Laramide events.

Large regions of anomalously shallow ocean floor, 100,000 km^2 or more in area, are found in the Pacific, Indian and Caribbean ocean basins. These regions were formed between the late Jurassic and the late Cretaceous and lie at depths several kilometres shallower than contemporaneous normal oceanic crust[7]. The plateaus have been interpreted as composed of abnormally thick and more buoyant oceanic crust[8].

Regions such as the Ontong Java and Manihiki plateaus, the Hess Rise and the eastern portion of mid-Pacific mountains seem to have formed by areally extensive, on ridge-axis, volcanic events between 120 and 90 Myr BP (Barremian to Turonian)[9–11]. Conversely, the western portion of the mid-Pacific mountains and the present oceanic crust of the Caribbean area seem to be a result of thickening of normal oceanic crust by voluminous off ridge-axis volcanic activity[9–11]. The shallowness of the Caribbean ocean floor is related to the late Cretaceous intrusion of large volumes of sills that thickened and made more buoyant the oceanic crust of this area[8]. Plateau regions that are generated by intense on ridge-axis volcanism may form symmetrical objects or 'twins' on two plates[12].

Regions of unusually buoyant ocean floor are difficult to subduct[8,13–15]. Therefore, the subduction or attempted subduction of oceanic plateaus or linear aseismic ridges can be expected to have a profound effect on the internal tectonics of the associated arc[8,13–15].

Because oceanic plateau regions are composed of crust that has a buoyancy between normal oceanic crust and continental crust, the subduction of such plateaus will cause the subducted lithospheric plate to ride higher in the asthenosphere, considerably lowering the angle of subduction[15]. For example, beneath the South American Cordillera the subducted oceanic plate has been segmented along strike, resulting in alternating sections of high- and low-angle subduction of the same plate[16]. Segments characterized by low-angle subduction correspond with portions of the cordilleran belt that lack arc-related Quaternary volcanism. Above the segments of high-angle subduction, Quaternary volcanism is abundant. An extremely important observation is that the segments of low-angle subduction beneath Peru and central Chile correspond with places where the eastern extensions of the Nazca and Juan Fernandez aseismic ridges have been subducted[16–19]. This suggests that the thickened and more buoyant oceanic crust of both the Nazca and Juan Fernandez aseismic ridges causes the descending oceanic plate to ride higher in the asthenosphere, resulting in a lower angle of subduction[17,19]. Enhanced seismic activity in the upper 50 km of the overriding South American plate[20] and extensive high- and low-angle thrust faulting in the foreland regions of Peru and Ecuador up to 700 km east of the South American trench[19,21,22] are associated with these sections of low-angle subduction.

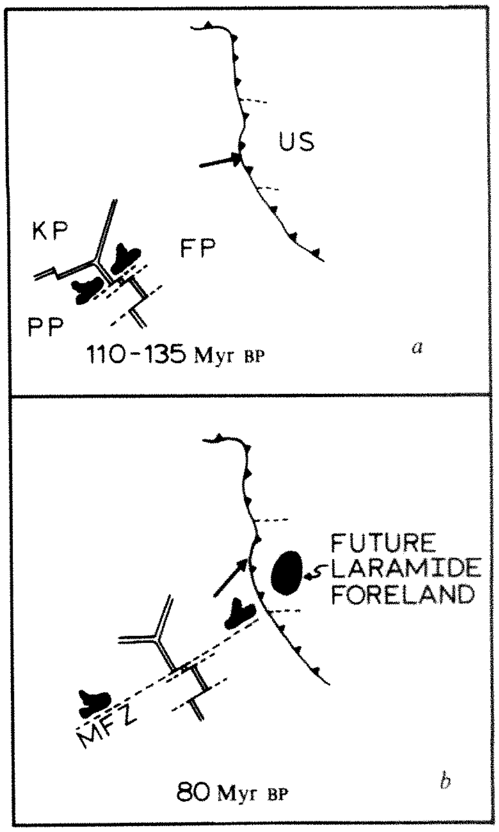


Fig. 2 Cretaceous reconstructions of the Kula, Farallon, Pacific and North American plates (modified from ref. 33): *a*, The formation of the Hess Rise and Hess Rise 'twin' by an unusually intense on axis volcanic event during the early Cretaceous; *b*, the movement of the Hess Rise 'twin' towards the future Laramide foreland during medial and late Cretaceous. Black arrows show the relative motion between the North American and Farallon plates[33] (KP, Kula Plate; FP, Farallon Plate; PP, Pacific Plate; MFZ, Mendicino Fracture Zone).

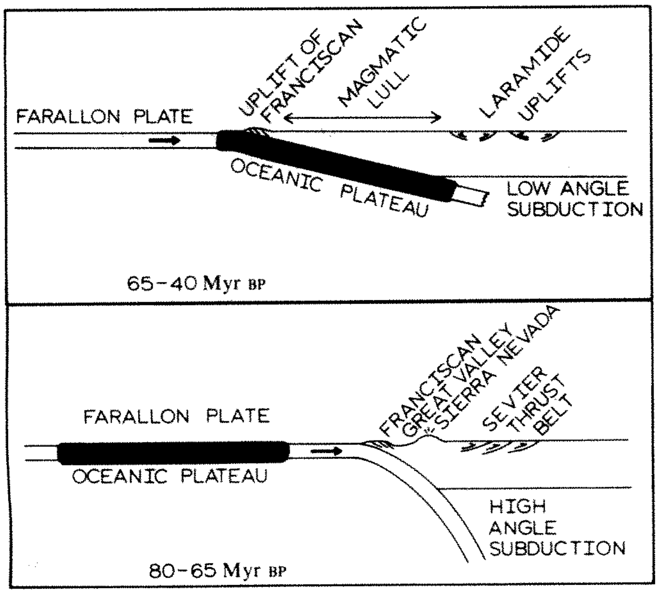


Fig. 1 Schematic cross-sections of the western US cordillera illustrating the progressive approach of an oceanic plateau towards the western US during the medial Cretaceous times and eventual subduction of the plateau during the Laramide orogeny.

The Laramide orogeny was characterized by the development of basement-cored overthrusts and adjacent synkinematic depositional basins in the foreland of the western US between northern Montana and northern New Mexico[23]. Deformation began ~70 Myr BP and continued until ~40 Myr BP (refs 24, 25). To the north and south of the Laramide foreland, Sevier-type retroarc thrusting continued[4,25]. The Laramide orogeny was a non-collisional, compressional event that resulted from the collapse and telescoping of the entire foreland region between Montana and New Mexico[26]. Faults that bound the ranges seem, mostly, to be low-angle thrusts that do not steepen at depth, indicating that horizontal compression and collapse of the foreland region was responsible for development of Laramide structures[22,27]. To the west of the Laramide foreland, arc volcanism ceased along the Sierra Nevada[28] and migrated inland, implying a rapid shallowing in the angle of subduction of the Farallon plate beneath the western US[29,30].

All these events may be interpreted as being the result of subduction of an oceanic plateau adjacent to this region. Subduction of a plateau of approximately the same dimensions as the Laramide foreland would result in flattening of the subduction angle, cessation of arc volcanism and increased compressive deformation, resulting in collapse in the foreland region (Figs 1 and 2). Therefore, tectonic patterns that characterized the Laramide orogeny such as cessation of arc volcanism and intense foreland deformation may be explained by inferring the subduction of an oceanic plateau adjacent to the Laramide foreland in the late Cretaceous/early Palaeogene, analogous with the situation presently found in South America.

Although it cannot be directly proved that such an oceanic plateau region ever existed or was subducted beneath the western US, the formation of numerous oceanic plateau regions in Cretaceous times, such as the Ontong Java and Manihiki plateaus, the mid-Pacific mountains and the present oceanic crust of the Caribbean, implies that these and perhaps other plateau regions lay within the Pacific Ocean during the late Mesozoic.

The Hess Rise is an oceanic plateau up to 800 km across lying just north of the westernmost extension of the Mendocino fracture zone and just south of the Great Magnetic Bight[31] (Fig. 2). The oldest sediments found on the Hess Rise are early Albian in age[10,32], implying that the Rise was formed sometime in the early Cretaceous. Watts *et al.*[11] and Vallier *et al.*[10] have presented evidence demonstrating that the Hess Rise formed by an intense on-axis volcanic event between 120 and 90 Myr BP (Barremian to Turonian). If this was the case, then symmetrical plateau twins may have been generated on two plates[12] (Fig. 2). That is, a Hess Rise twin could have formed on the Farallon plate symmetrically with the Hess Rise on the Pacific Plate.

Figure 2 shows reconstructed plate geometries for the Cretaceous[33] and the approximate positions of both the Hess Rise and postulated Hess Rise twin. Symmetrical spreading about the Pacific–Farallon ridge brings the Hess Rise twin into a position just opposite the future Laramide foreland by the late Cretaceous (Fig. 2). Continued convergence between the North American and Farallon plates at the rapid rate of ~ 14 cm yr^{-1} (ref. 33) would have resulted in subduction of a Hess Rise twin opposite the Laramide foreland during the late Cretaceous through the early Cenozoic. Thus, subduction of a Hess twin beneath the western US may have been the cause of the distinctive Laramide orogeny.

Indirect evidence within the Cordillera for the subduction of an oceanic plateau beneath the western US during the Laramide orogeny may exist. Because plateau regions lie at depths 2–3 km above oceanic crust of the same age, then during the subduction of an oceanic plateau, uplift would be expected along the entire fore-arc region as it rose to accommodate oceanic crust riding at a much higher elevation. Dickinson *et al.*[34] have documented such an event for the forearc region of California in the late Cretaceous that lifted the trench slope-break to a position near sea level resulting in widespread subaerial unconformities and filling of the fore-arc trough to the east to form a wide fore-arc platform area. Also, Winterer[35] has suggested that the Calera and Laytonville limestones found within the Franciscan melange may represent obducted slices of an oceanic plateau.

We thank J. F. Dewey, R. H. Pilger and W. R. Dickinson for preprints, and R. V. Ingersoll, R. H. Pilger, T. A. Cross and W. S. F. Kidd for reviews. Research at the State University of New York at Albany on oceanic plateaus has been partly supported by NSF grants 7614754 and 7803319.

Received 30 May; accepted 11 November 1980.

1. Woodward, L. A. *New Mex. geol. Soc. Spec. Publ.* **6,** 11 (1976).
2. Coney, P. J. *Am. J. Sci.* **272,** 603 (1972).
3. Dickinson, W. R. & Snyder, W. S. *Geol. soc. Am. Mem.* **151,** 355 (1978).
4. Burchfiel, B. C. & Davis, G. A. *Am. J. Sci.* **275A,** 363 (1975).
5. Molnar, P. & Atwater, T. *Earth. planet. Sci. Lett.* **41,** 330 (1978).
6. Dewey, J. F. *Geol. Ass. Can: Spec. Publ.* **20,** (in the press).
7. Sclater, J. G., Anderson, R. N. & Bell, M. L. *J. geophys. Res.* **76,** 7888, (1971).
8. Burke, K., Fox, P. J. & Sengör, A. M. C. *J. geophys. Res.* **83,** 3949 (1978).
9. Winterer, E. L. *Am. geophys. Un. Monogr.* No. 19, 269 (1976).
10. Vallier, T. L. *et al. Nature* **286,** 48 (1980).
11. Watts, A. B., Bodine, J. H. & Ribe, N. M. *Nature* **283,** 532 (1980).
12. Sharman, G. F. *EOS* **60,** 949 (1979).
13. Vogt, P. R., Lowrie, A., Bracey, D. R. & Hey, R. N. *Geol. Soc. Am. Spec. Pap.* 172 (1976).
14. Hughes, G. W. & Turner, C. C. *Bull. geol. Soc. Am.* **88,** 412 (1977).
15. Kelleher, J. & McCann, W. *Am. geophys. Un. Maurice Ewing Ser.* **1,** 115 (1977).
16. Isacks, B. L. & Barazangi, M. *Am. geophys. Un. Maurice Ewing Ser.* **1,** 99 (1977).
17. Pilger, R. H. *EOS* **58,** 1232 (1977).
18. Aleman Ramirez, A. M. & Pilger, R. H. *EOS* **59,** 383 (1978).
19. Pilger, R. H. *Bull. geol. Soc. Am.* (in the press).
20. Barazangi, M. & Isacks, B. L. *Geology* **4,** 686 (1976).
21. Burchfiel, B. C. & Davis, G. A. *Nature* **260,** 693 (1976).
22. Brewer, J. A., Smithson, S. B., Oliver, J. E., Kaufman, S. & Brown, L. D. *Tectonophysics* **62,** 165 (1980).
23. Hamilton, W. *Soc. Econ. Paleont. Miner. Pacific Coast Paleogeogr. Symp.* **2,** 33 (1978).
24. Tweto, O. *Geol. Soc. Am. Mem.* **144,** 1 (1975).
25. Armstrong, R. L. *Nature* **247,** 348 (1974).
26. Coney, P. J. *New Mex. geol. Soc. Spec. Publ.* **6,** 5 (1976).
27. Smithson, S. B., Brewer, J., Kaufman, S., Oliver, J. & Hurich, C. *Geology* **6,** 648 (1978).
28. Snyder, W. S., Dickinson, W. R. & Silberman, M. L. *Earth planet. Sci. Lett.* **32,** 91 (1976).
29. Coney, P. J. & Reynolds, S. J. *Nature* **270,** 403 (1977).
30. Cross, T. A. & Pilger, R. H. Jr *Am. J. Sci.* **278,** 865 (1978).
31. Chase, T. E., Menard, H. W. & Mammerickx, J. *I.M.R. Tech. Rep. Ser. TR*-17 (1971).
32. Scientific Party *Init. Rep. DSDP* Leg **32,** 233 (1975).
33. Coney, P. J. *Geol. Soc. Am. Mem.* **152,** 33 (1978).
34. Dickinson, W. R., Ingersoll, R. V. & Graham, S. A. *Bull. geol. Soc. Am.* **90,** 1458 (1979).
35. Winterer, E. L. *Texas A & M Univ. Symp. on Oceanic Plateaus,* 15 (1979).

Old model Nd ages in Namibian Pan-African rocks

C. J. Hawkesworth*, J. D. Kramers† & R. McG. Miller‡

* Department of Earth Sciences, The Open University,
Walton Hall, Milton Keynes MK7 6AA, UK
† Department of Earth Sciences, The University,
Leeds LS2 9JT, UK
‡ Geological Survey, PO Box 2168, Windhoek, 9100, Namibia

The Nd-isotope analytical technique is a powerful tool for studying many geological processes: particularly the evolution of the upper mantle as seen through the isotope and trace element geochemistry of mantle-derived volcanic rocks[1–3], and magmatic processes along destructive plate margins[4,5]. However, previous studies of continental crust[6–8] have been largely restricted to Archaean examples and concerned primarily with dating their time of derivation from the upper mantle. In the present study we investigated the Nd- and Sr-isotope characteristics of rocks involved in a relatively young (800–450 Myr) orogenic event. Such events represent critical stages in the evolution of most of the Earth's continental crust—new material is added from the mantle, and pre-existing crust is remobilized by erosion and sedimentation, deformation and magmatic activity. Nd- and Sr-isotopes were used to outline the range in age and geochemical characteristics of rock sequences and provinces within both the upper mantle and the pre-existing crust which were sampled during the orogeny, and hence provided the major components of what is now a stable segment of continental crust.

The Damara high-temperature belt of Namibia (Fig. 1*a*), formed during the widespread orogenic event first categorized as Pan African by Kennedy[9] and it was chosen because it is arguably the best exposed orogenic belt of its age and has been studied in detail both geologically and geochemically (refs 10–13 and unpublished results). Some aspects of the stratigraphy and in particular the age and regional significance of the earlier tectonic events are still poorly understood, but an outline is presented in Fig. 1*b*.

Radiometric ages on pre-Damara rocks range from 1,100 Myr along the southern margin[14,15] to 1,700–2,000 Myr in the centre of the belt[16] and near the northern margin[17]. Within the Damara succession, the first major magmatic event is recorded in a suite of large ionic lithophile (LIL)-element-enriched lavas (Naauwpoort volcanics) and associated syenites and carbonatites exposed along the northern margin. U/Pb zircon results are variable, but they seem to cluster around an age of 800 Myr (ref. 18). In the north, Damara sediments reflect a well developed shelf facies, whereas in the centre and south, clastic sediments with only minor carbonates are followed by a thick flysch-type succession (Kuiseb schists). Within the flysch there is a thin band of basic rock, the Matchless amphibolite, which has been interpreted both as a fragment of tectonically emplaced ocean floor material[19], and as just one of several basic volcanic units within the flysch trough[10]. Whichever interpretation is correct an Rb/Sr whole-rock age of 765 ± 37 Myr

0028-0836/81/040278—05$01.00

from this Matchless amphibolite (A. Kröner, personal communication) is probably a minimum age for most of the sediments still preserved in the southern zone.

Granitic rocks, varying from diorites to granites and highly potassic alaskites, dominate the centre of the Damara belt (Fig. 1). Some, which intrude Damara sediments and have undergone extensive regional deformation, yield ages of 700–750 Myr (A. Kröner, personal communication). They are loosely termed 'early' granites and their relationship to later events is not understood. 'Late' granites, by contrast (Fig. 1*b*), all post-date the major regional deformation and many have been studied previously (refs 12, 20, 21 and unpublished results). Initial Sr-isotope ratios vary from 0.705 to 0.759 (unpublished results) and many of the younger granites, which tend to have the higher initial ^{87}Sr/^{86}Sr ratios were probably emplaced during late uplift in the centre of the belt.

The Damara orogenic cycle may therefore be broadly subdivided into four phases: (1) clastic sedimentation and LIL-element-enriched magmatism; (2) basic magmatism (possibly ocean floor) and predominantly flysch-type sedimentation; (3) early granites and regional tectonism; and (4) late, largely post-tectonic granites with associated high-temperature metamorphism. Our samples included volcanic and sedimentary rocks and a range of lithologies from the late granites (Fig. 1).

^{143}Nd/^{144}Nd ratios and the concentration of eight rare earth elements (REEs) were determined using techniques described previously[4] and the results are presented in Tables 1 and 2. Except for K211, all samples were selected from suites which plot on whole-rock Rb–Sr isochrons (unpublished results). Their initial ^{87}Sr/^{86}Sr ratios are thus representative of the particular rock unit. Initial Nd-isotope ratios vary from 0.51196 in the Matchless amphibolite to 0.51117 in the K_2O-rich Rössing alaskite. Figure 2 shows chondrite-normalized REE patterns for selected samples.

The carbonatite (28/32) from the Naauwpoort magmatic suite contains extremely high REE concentrations similar to those reported from carbonatites elsewhere[22]. Although we are unaware of other carbonatites with a similar 'concave-up' distribution (Fig. 2), such patterns are predicted by experimental work on the distribution of REEs between common mantle minerals and a CO_2-rich vapour phase[23].

The samples of Matchless amphibolite (AM 62, 63) exhibit higher Sm/Nd and ^{143}Nd/^{144}Nd ratios than the chondritic uniform reservoir (CHUR[2], presumed to be representative of the bulk Earth) at the time of their formation. This indicates that their source region had been relatively depleted in light REEs for a considerable time and, because such depletion is characteristic of recent mid-oceanic ridge basalts, these results are in keeping with the interpretation that the Matchless amphibolite represents a fragment of oceanic crust[19]. The Gariep Belt is another possible ophiolite sequence, probably of similar age to the Matchless, and whereas the dolerite sample (K211) is enriched in light REEs, it exhibits the same relatively high initial ^{143}Nd/^{144}Nd ratio (Table 1).

Kh 24 (Kuiseb schist) is a metamorphosed flysch-like sediment which was probably deposited 700–800 Myr ago (Fig. 1). Rb/Sr whole-rock results scatter about an 'errorchron' corresponding to an age of 548 ± 56 Myr, which has been interpreted as the time of widespread regional metamorphism (unpublished results). The REEs exhibit a negative Eu anomaly which is typical of post-Archaean sediments[24] and a Sm/Nd ratio in the range suggested for average continental crust[25] (0.17–0.21).

The Damara 'granites' sampled are from massive stocks of diorite and granite and from small bodies of K-rich alaskite. The diorite (RM586) and two granites (RM662 and SM2) range in age from 560 to 490 Myr, in initial ^{87}Sr/^{86}Sr from 0.7048 to 0.709, and in initial ^{143}Nd/^{144}Nd from 0.51186 to 0.51146. The diorite, with its initial Sr- and Nd-isotope ratios similar to CHUR 550 Myr ago, probably originated in the upper mantle although its generation must have involved considerable fractionation of the light REEs (Fig. 2). The granites have higher

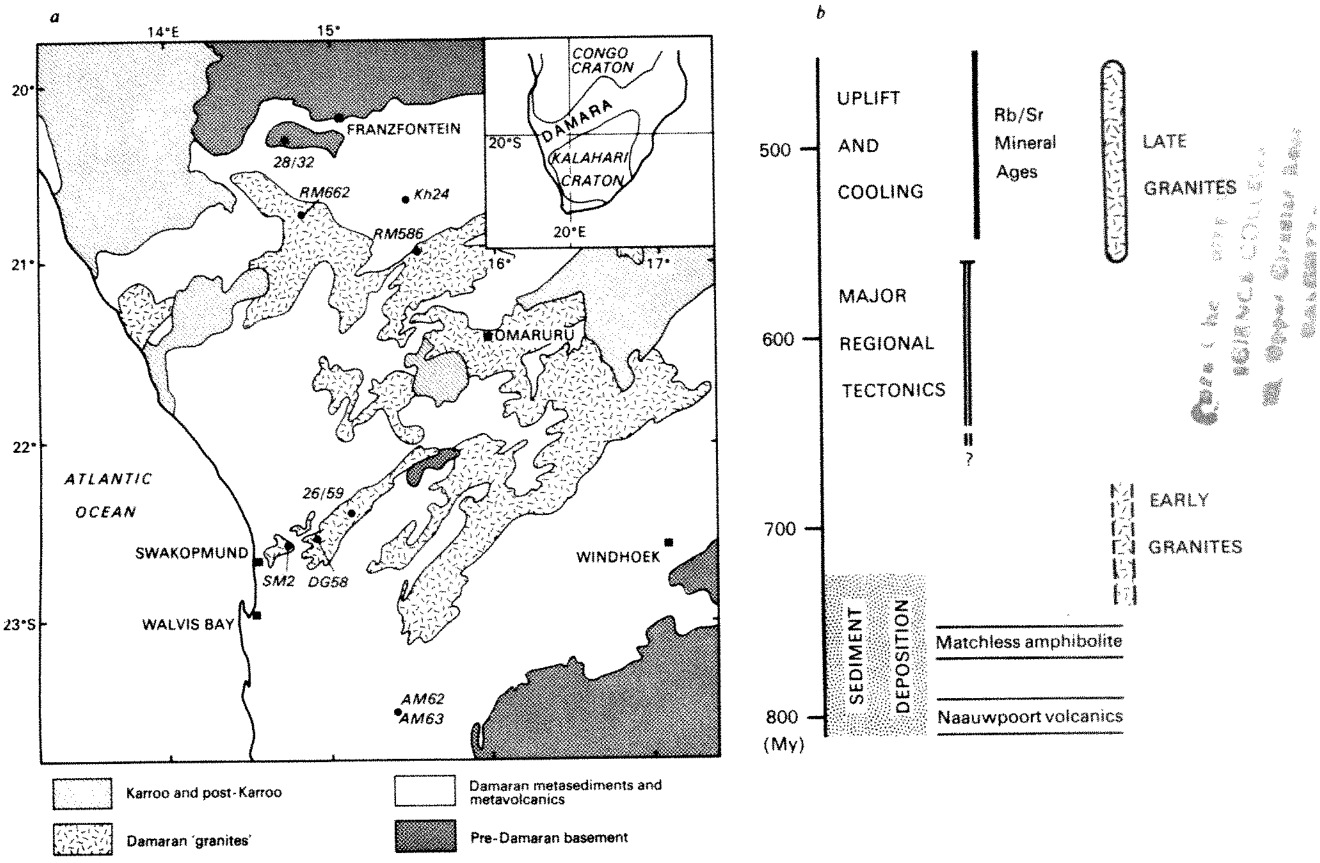


Fig. 1 *a*, Geological map of the Damara high-temperature belt, Namibia, illustrating sample localities. *b*, Summary of the major geological events recognized within the central Damara succession.

Table 1 Nd-isotope results

Rock no.	Lithology	*a* Age (Myr)	*b* $^{143}Nd/^{144}Nd_m$	*c* Sm/Nd	*d* $^{143}Nd/^{144}Nd_0$	*e* T^{Nd}_{CHUR} (Ga)	*f* $^{87}Sr/^{86}Sr_0$
AM62	Matchless amphibolite	765±37	0.513124±34	0.366	0.51196	2.5	0.70566±7
AM63	Matchless amphibolite	765±37	0.513093±18	0.365	0.51194	2.4	0.70566±7
K211	Gariep dolerite	~765	0.512440±26	0.143	0.51196	0.3	—
28/32	Lofdal carbonatite	~800	0.512201±20	0.181	0.51157	0.8	0.7032±3
Kh24	Kuiseb schist	548±56	0.512219±20	0.209	0.51171	1.0	0.715±2
RM586	Salem diorite	~550	0.512330±30	0.178	0.51186	0.6	0.7048±6
SM2	Salem granite	563±63	0.512153±28	0.189	0.51168	1.0	0.707±4
RM662	Sorris Sorris Granite	495±15	0.511795±16	0.144	0.51146	1.2	0.709±1
26/59–155.5m	Valencia alaskite	484±18	0.511735±24	0.219	0.51126	2.3	0.724±2
DG58	Rössing alaskite	458±8	0.511588±28	0.204	0.51117	2.3	0.759±1

a, Rb/Sr whole rock ages (unpublished results and A. Kröner, personal communication), except for the Lofdal carbonatite (U/Pb zircon[18]) and Gariep dolerite (assumed to be of the same age as the Matchless amphibolite).

b, Measured $^{143}Nd/^{144}Nd$ ratios. Normalized to $^{146}Nd/^{144}Nd = 0.7219$. Quoted errors are two standard errors on the mean. $^{143}Nd/^{144}Nd$, BCR-1 = 0.51266±2.

c, From Table 2, except AM63 (R.M., unpublished results).

d, Initial $^{143}Nd/^{144}Nd$ normalized to BCR-1 = 0.51262.

e, $$T^{Nd}_{CHUR} = \frac{1}{\lambda}\ln\left[\frac{^{143}Nd/^{144}Nd_{CHUR} - {}^{143}Nd/^{144}Nd_{sample}}{^{147}Sm/^{144}Nd_{CHUR} - {}^{147}Sm/^{144}Nd_{sample}} + 1\right]$$

where $\lambda = 6.54 \times 10^{-12}\ yr^{-1}$, $^{143}Nd/^{144}Nd_{CHUR} = 0.51262$, $^{147}Sm/^{144}Nd_{CHUR} = 0.1935$. The T^{Nd}_{CHUR} age for K211 is meaningless as its Sm/Nd ratio is probably much less than that in its source rock, whereas those for 26/59 and DG58 are probably slightly too old (see text).

f, Initial $^{87}Sr/^{86}Sr$ ratios (unpublished results and A. Kröner, personal communication).

initial $^{87}Sr/^{86}Sr$ and lower initial $^{143}Nd/^{144}Nd$ ratios and thus probably contain material with a crustal prehistory. Their much larger Eu anomalies (Fig. 2) show that their generation involved either separation from a plagioclase-bearing residue, or plagioclase fractionation. Alaskites in contrast were produced by partial melting at high levels in the crust, sometimes apparently *in situ*[26]. Samples 26/59 and DG58 are from 480- and 460-Myr old bodies with unusually high initial $^{87}Sr/^{86}Sr$ (0.724, 0.759) (ref. 21 and unpublished results) and low initial $^{143}Nd/^{144}Nd$ (0.51126, 0.51117) ratios. Note the low concentrations of Sr and the medium- and light-REE, especially in the samples from Rössing (DG58 and 59) and the 'concave-up' distribution patterns (Fig. 2). The low Sr concentrations are consistent with a small degree of partial melting in equilibrium with residual plagioclase. The REE distribution patterns could reflect the presence of minor phases in the residuum, or the scavenging of REE with U in the formation of the adjacent U mineralization. Rössing, for example, is the site of a major uranium deposit and monazite, xenotime and sphene (all minerals with very high crystal–liquid partition coefficients for medium- and light-REE)[27] are commonly found in association with the uranium mineralization (refs 4, 28, and Marlow, personal communication).

The initial $^{143}Nd/^{144}Nd$ ratios of these Pan African rocks are plotted against time in Fig. 3, with the evolution of the chondritic reservoir (CHUR[2]) shown for reference. The possible ocean-floor volcanics (amphibolites and dolerite) and the carbonatite and diorite have initial Nd-isotope ratios respectively higher than and similar to that of CHUR at the time of their formation, whereas the late granites and alaskites have progressively lower initial $^{143}Nd/^{144}Nd$ ratios. Moreover, as there is a broad inverse correlation between initial Sr- and Nd-isotopes (expressed as ε_{Nd} and ε_{Sr} respectively[29], Fig. 4), the lower initial $^{143}Nd/^{144}Nd$ ratios in the younger rocks is matched by their progressively higher $^{87}Sr/^{86}Sr$ ratios.

T^{Nd}_{CHUR} ages[8] are estimates of how long a rock (and its source region) have had different Sm/Nd, and hence $^{143}Nd/^{144}Nd$ ratios to CHUR. They denote the time at which the Nd-isotope evolution curve of a system with particular present day $^{143}Nd/^{144}Nd$ and Sm/Nd ratios intersects that of CHUR (see Fig. 3). T^{Nd}_{CHUR} ages are model ages because it is assumed that a simple two-stage history applies. The first stage is in a (mantle) province with the isotope and trace element characteristics of CHUR, and the second stage reflects the $^{143}Nd/^{144}Nd$ and Sm/Nd ratios in the rocks at present. If such a model is inapplicable as, for example, when light REE-enriched crustal rocks are derived from depleted mantle (higher $^{143}Nd/^{144}Nd$ than CHUR), the time of separation from the mantle may be

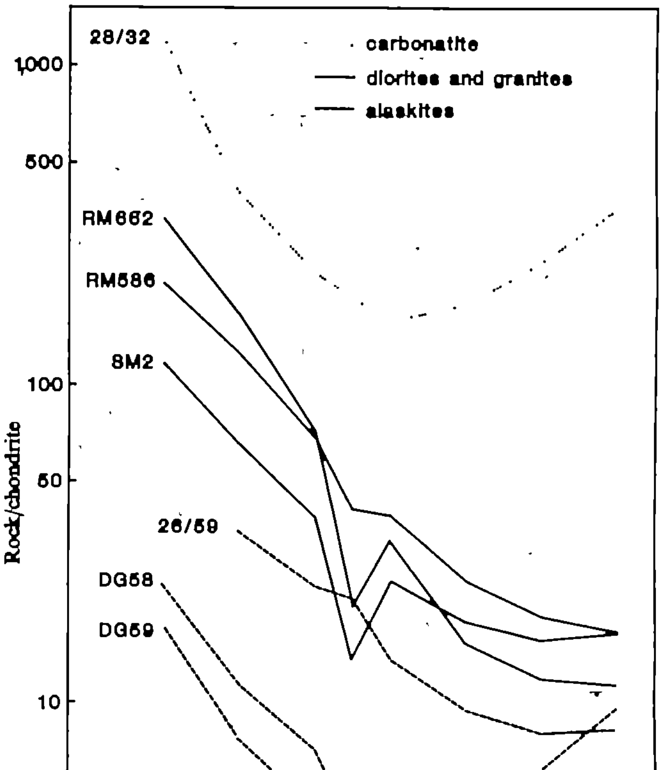


Fig. 2 Chondrite-normalized REE distribution patterns for selected Damara rocks.

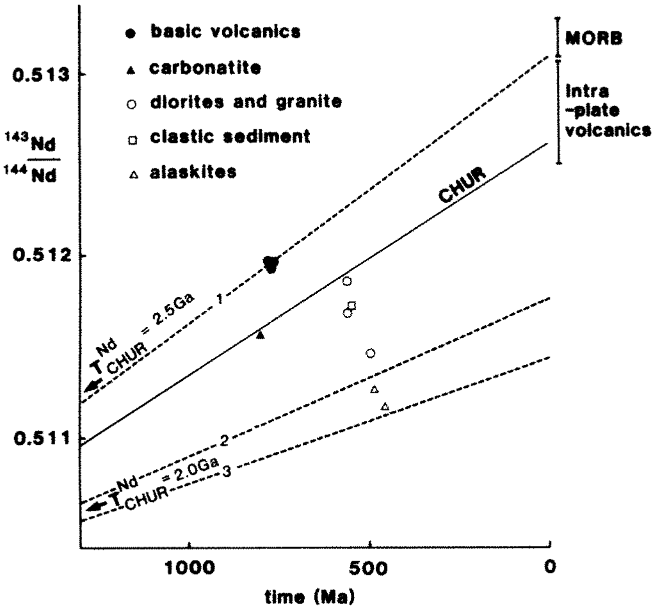


Fig. 3 Variations in initial ^{143}Nd/^{144}Nd ratios in the Damara rocks. CHUR (chondrite uniform reservoir), ^{147}Sm/^{144}Nd = 0.1935, present day ^{143}Nd/^{144}Nd = 0.51262. Evolution paths; 1, Sm/Nd = 0.366 (as in the samples of Matchless amphibolite); 2 and 3, Sm/Nd = 0.214 and 0.173 respectively (that is the range in average continental crust[25]). The samples between evolution paths 2 and CHUR have T^{Nd}_{CHUR} ages in the range 550–1,200 Myr (Table 1).

older than that implied by their T^{Nd}_{CHUR} age. Nonetheless, for mantle-derived volcanics, such model Nd ages reflect how long chemical variations have persisted in that part of the upper mantle, and for crustal-derived rocks (be they magmatic or sedimentary) they provide a useful estimate of the age of their source region. In all cases they require that the Sm/Nd and ^{143}Nd/^{144}Nd ratios of the rocks are similar to those in their source areas; inevitably this assumption is more valid for some rock types than for others.

The basic volcanic rocks probably reflect 10–20% partial melting of spinel peridotite and thus their Sm/Nd ratios are likely to be similar to those in their source regions[3]. Thus the evolution path of material with Sm/Nd = 0.366 (that is, as in AM62) has been plotted in Fig. 3; T^{Nd}_{CHUR} equals 2,500 Myr, and its present day ^{143}Nd/^{144}Nd is similar to those in recent MOR basalts. This suggests that these rocks were derived from mantle which had been depeleted in light REE since the Archaean and supports the interpretation that both the Matchless and Gariep Belts represent fragments of oceanic crust[19]. The anomalously high ε_{Sr} values of these samples (Fig. 4) probably reflects contamination, either with seawater on the ocean floor[30], or during post-emplacement metamorphism.

The low concentrations of REEs in the alaskites and 'concave-up' distribution patterns (Fig. 2) indicate that they have probably been modified—perhaps by processes associated with the uranium mineralization. Thus, rather than assuming that the Sm/Nd ratios of the individual alaskites were the same as their source rocks, we have plotted the evolution of material which was derived from a chondritic reservoir (CHUR) 2,000 Myr ago (that is, the oldest age reported from basement rocks in the Damara[16]) and had Sm/Nd ratios in the range for average continental crust (0.173–0.214, see ref. 25 for discussion). Its evolution path in Fig. 3 encompasses the low initial Nd-isotope ratios of the alaskites which strongly suggests that they were derived from crustal material which had been out of the mantle since 2,000 Myr.

The initial ^{143}Nd/^{144}Nd ratios of the granites (SM2, RM662) and the metasediment (Kh24) are intermediate between those of the alaskites and CHUR at the time of their formation (Fig. 3). Apart from equilibrium with minor phases (as presumably happened in the case of the alaskites), equilibration with a large amount of cumulus or residual amphibole is the most likely mechanism capable of altering significantly the Sm/Nd ratios of a crustal melt[31]. This would be reflected in a strong depletion in heavy REEs and, because this is not observed in any of these samples, we conclude that the Sm/Nd ratios of the granites SM2 and RM662 are likely to be similar to those in their source rocks. Sample Kh24, being a metasediment, is also likely to have an Sm/Nd ratio similar to that in its source terrain[8]. The T^{Nd}_{CHUR} ages for these two granites and the metasediment range from 1,000 to 1,200 Myr indicating that, unlike the alaskites, they were not derived from 2,000-Myr old crustal material alone. Two possibilities remain: the granites (and the metasediment) could simply have been derived from younger crustal sources (perhaps ~1,100-Myr old) or they could contain variable amounts of both mantle-derived and 2,000-Myr crustal material. For the granites this might reflect contamination of mantle-derived magmas within the crust, whereas the flysch-type sediment could contain a mixture of young volcanic and older continental detritus. At present there are insufficient data to choose between these two hypotheses, however: (1) significant quantities of new crustal material was generated 1,100 Myr ago along the present-day southern margin of the Damara[14,15], and crustal rocks of this age may also occur at depth elsewhere within the belt; (2) where contamination of magmas with continental crust has been convincingly demonstrated, it tends to result in a flat-lying trend between Nd- and Sr-isotopes[5]. This is not observed for the Damara diorite and granites (Fig. 4).

The inverse correlation of ε_{Sr} and ε_{Nd} observed in the crustal rocks studied is interesting. The inverse correlation in mantle-derived volcanics (also shown in Fig. 4) is generally accepted as an expression of coherent fractionation of Rb/Sr and Nd/Sm ratios during melting in the mantle. However, magmatic and sedimentary processes, particularly in the upper crust, fractionate Rb/Sr much more readily than Nd/Sm. Thus sediments and K_2O-rich granitic material develop with time relatively high ^{87}Sr/^{86}Sr ratios, resulting in a distinctive, curved inverse correlation between ε_{Nd} and ε_{Sr} at higher levels in the continental crust (Fig. 4). Furthermore, although magmatism in this orogenic event has sampled material from the upper mantle through to the upper continental crust, it does not seem to have sampled

Table 2 Trace elements (p.p.m.)

Rock no.	Rb*	Sr*	Rb/Sr	Ce	Nd	Sm	Eu	Gd	Dy	Er	Yb
AM62	—	—	—	8.44	7.85	2.87	1.01	4.16	5.15	3.41	3.17
K211	105	153	0.686	374	88	12.6	2.98	9.69	9.50	5.60	5.61
28/32	9.4	6,660	0.0014	1,062	249	45.1	14.2	43.9	61.7	54.1	77.9
Kh24	78	139	0.561	101	46.9	9.80	1.78	7.95	8.42	5.13	5.21
RM586	71	1,045	0.68	180	79	14.1	3.10	10.9	8.25	4.17	3.70
SM2	183	126	1.45	103	41.2	7.79	1.06	6.72	6.13	3.50	3.60
RM662	213	205	1.04	293	101	14.5	1.59	8.81	5.21	2.65	2.46
26/59	250	123	2.03	—	21.6	4.73	1.65	3.76	3.21	1.78	1.83
DG58	334	28	11.9	20.3	7.04	1.44	0.30	1.28	1.78	1.37	2.11
DG59	315	18	17.5	15.2	4.78	0.94	0.31	0.93	1.16	0.85	1.28

* Unpublished results.

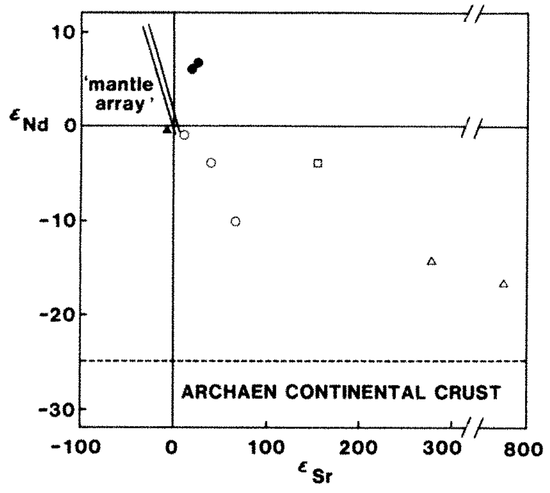


Fig. 4 ε_{Nd} against ε_{Sr} (ref. 2) for Damara rocks (key as for Fig. 3)

$$\varepsilon_{Nd}=\left[\frac{^{143}\mathrm{Nd}/^{144}\mathrm{Nd(initial)}}{^{143}\mathrm{Nd}/^{144}\mathrm{Nd}_{\mathrm{CHUR}(t)}}-1\right]\times 10^4$$

where ^{143}Nd/^{144}Nd$_{\mathrm{CHUR}(t)}$ is the ^{143}Nd/^{144}Nd ratio of a chondritic uniform reservoir at the time of formation of the rock in question. ε_{Sr} is similarly defined using the relevant ^{87}Sr/^{86}Sr ratios[2]. ε_{Nd} and ε_{Sr} calculated using the ages and initial Nd- and Sr-isotope ratios in Table 1, but note that the age of the sediment (Kh24) is believed to reflect regional metamorphism (unpublished results).

any old granulite facies rocks ($-\varepsilon_{Nd}$ and $-\varepsilon_{Sr}$, Fig. 4, refs 1, 7). Such rocks may not be ubiquitous in the lower crust, but it is also possible that such LIL-element depleted material is not readily mobilized.

We conclude that Pan African rocks in Namibia were derived from upper-mantle and continental sources which at the time of the Damara orogeny had already been respectively depleted and enriched in light rare earths for up to 2,000 Myr. High initial ^{143}Nd/^{144}Nd ratios on basic volcanic rocks support suggestions that they represent fragments of oceanic crust[18], and imply that their source regions had been depleted in light REEs since the Archaean. $T^{\mathrm{Nd}}_{\mathrm{CHUR}}$ ages of the granitic rocks range from 550 to 2,000 Myr. The older ages are from alaskites which, on field evidence, were generated at higher levels in the crust. The lower crust seems to contain younger (?1,100 Myr) crustal material. More generally this evidence for the remobilization of 2,000-Myr old crust highlights an important difference between high-temperature intracratonic Pan-African belts such as the Damara (Fig. 1) and those in north-east Africa which appear to be less deeply eroded, contain no evidence for significantly older continental crust, and are believed to reflect crustal accretion along convergent plate boundaries[32]. The Damara granitic rocks exhibit an inverse correlation between initial Nd- and Sr-isotopes (Fig. 4), although the alaskites and the metasediment are displaced to comparatively high ^{87}Sr/^{86}Sr ratios consistent with a relative increase in Rb/Sr at higher levels in the continental crust.

We thank M. P. Coward, C. J. Hartnady, J. Barnes, K. Downing and A. G. Marlow for discussions and A. R. Gledhill, J. C. Roddick and A. Kröner for unpublished results and comments, and also Professor Nicolaysen for his constructive review. Analyses were carried out at Leeds where isotope and REE research is funded by NERC and the Royal Society.

Received 9 June; accepted 20 November 1980.

1. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *Phil. Trans. R. Soc.* (in the press).
2. DePaolo, D. P. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **43,** 615–627 (1979).
3. Hawkesworth, C. J., Norry, M. J., Roddick, J. C. & Vollmer, R. *Nature* **280,** 28–31 (1979).
4. Hawkesworth, C. *et al. Earth planet. Sci. Lett.* **42,** 45–57 (1979).
5. Hawkesworth, C. J. in *Origin of Granite Batholiths: Geochemical Evidence*, 76–89 (Shiva, Kent, 1979).
6. Jacobsen, G. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **41,** 245–253 (1978).
7. Hamilton, P. J., Evensen, N. M., O'Nions, R. K. & Tarney, J. *Nature* **277,** 25–28 (1979).
8. McCulloch, M. T. & Wasserburg, G. J. *Science* **200,** 1003–1011 (1978).
9. Kennedy, W. Q. *8th A. Rep. Res. Inst. Afr. Geol.* **48** (University of Leeds, 1964).
10. Kröner, A. *Tectonophysics* **40,** 101–135 (1977).
11. Martin, H. & Porada, H. *Precamb. Res.* **5,** 311–338 (1977).
12. Blaxland, A., Gohn, E., Haack, U. & Hoffer, E. *Neues. Jb. Miner. Mh.* **11,** 498–508 (1979).
13. Coward, M. (in preparation).
14. Watters, B. R. *Nature* **259,** 471–473 (1976).
15. Malling, S. *A. Rep. Precamb. Res. Unit* Vol. 15, 183–193 (University of Cape Town, 1978).
16. Jacob, R. E., Kröner, A. & Burger, A. J. *Geol. Rdsch.* **67,** 706–718 (1978).
17. Burger, A. J., Clifford, T. N. & Miller, R. McG. *Precamb. Res.* **3,** 415–431 (1976).
18. Burger, A. J. & Walraven, F. *Annls. Geol. Surv. S. Afr.* **11,** 323–326 (1978).
19. Burke, K., Dewey, J. F. & Kidd, W. S. F. *Tectonophysics* **40,** 69–99 (1977).
20. Miller, R. McG. *Geol. Surv. S. Afr. Mem.* **64** (1973).
21. Kröner, A. & Hawkesworth, C. J. *Rep. Res. Inst. Afr. Geol.* Vol. 20, 14–16 (University of Leeds, 1977).
22. Mitchell, R. H. & Brunfelt, A. O. *Contr. Miner. Petrol.* **52,** 247–259 (1975).
23. Wendlandt, R. F. & Harrison, W. J. *Yb Carnegie Inst. Wash.* **77,** 695–703 (1978).
24. Nance, W. B. & Taylor, S. R. *Geochim. cosmochim. Acta* **40,** 1539–1551 (1976).
25. Shaw, D. M., Dostal, J. & Keays, R. R. *Geochim. cosmochim. Acta* **40,** 73–83 (1976).
26. Sawyer, E. W. thesis Univ. Cape Town (1978).
27. Condie, K. C. *Chem. Geol.* **21,** 131–149 (1978).
28. Berning, J., Cooke, R., Hiemstra, S. A. & Hoffman, U. *Econ. Geol.* **71,** 351–368 (1976).
29. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **4,** 465–468 (1977).
30. O'Nions, R. K., Carter, S. R., Cohen, R. S., Evensen, N. M. & Hamilton, P. J. *Nature* **273,** 435–438 (1978).
31. Arth, J. G. & Hanson, G. N. *Geochim. cosmochim. Acta* **39,** 325–362 (1975).
32. Gass, I. G. *J. geol. Soc. Lond.* **134,** 129–138 (1977).

A model for Plinian eruptions of Vesuvius

M. F. Sheridan*, F. Barberi, M. Rosi & R. Santacroce

Istituto di Mineralogia e Petrografia, via S. Maria 53,
56100 Pisa, Italy

The term 'Plinian' has been widely used[1–4] to describe continuous gas-blast eruptions of large magnitude a typical example[5], of which is the AD 79 eruption of Vesuvius which destroyed Pompei and the surrounding region. We develop a new model here for the AD 79 event that explains the complete Plinian eruptive episode including pyroclastic fall, pyroclastic flow, base surge, laharic and phreatic activity. This model has widespread implications with regard to volcanic hazard evaluation and geothermal exploration at Vesuvius and other volcanoes with similar patterns of activity, such as Mount St Helens.

Three representative sections from archaeological excavationsdemonstrate the principal characteristics of the entire eruptive sequence, assuming emplacement of pyroclastic flows, surges and lahars above the pumice bed, during the 30-h eruptive period. Data from more than 30 sections included in the new model show that factors such as distance from the vent, topography and wind velocity, control the depositional facies.

At the Pompei excavation 4.5-m of section (Fig. 2) includes a lower white pumice-fall layer that grades into an upper grey, pumice-fall bed overlain by interbedded surge and grey pumice-fall levels and an uppermost surge sequence. There is an upward increase in lithic xenoliths from 12 wt% for the white part to 20 wt% for the grey[5]. The character of the xenoliths also changes with a slight increase in fraction of carbonates and a strong increase in cognate cumulates in the grey level. The lower group of surge beds is of the dry (superheated steam) massive and sandwave bed type[9,10]. This group includes two prominent phreatic fall levels rich in scoria and calcareous fragments. The contact between the grey pumice and the first surge beds is a prominent wave form. The uppermost part of the surge sequence probably represents emplacement of very wet (condensed steam) phreatomagmatic products. These massive fine-ash beds are full of accretionary lapilli and may have the widespread dispersal characteristic of phreatoplinian deposits[11].

At the Oplonti excavation the section is 8.5-m thick (Fig. 2). The grey pumice-fall layer, which is much thicker than the white,

---

* Present address: Department of Geology, Arizona State University, Tempe, Arizona 85281.

0028-0836/81/040282—04$01.00

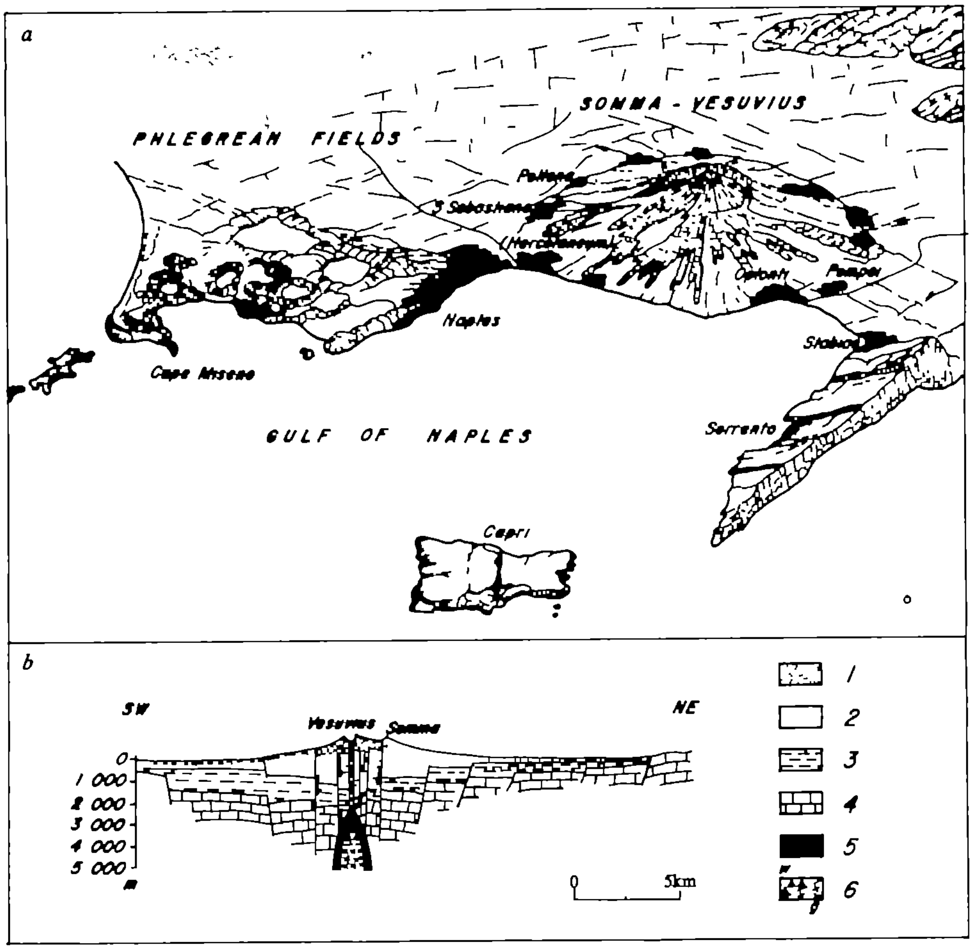


Fig. 1 Somma–Vesuvius Volcano. *a*, Location map. *b*, Schematic cross-section. Key: *1*, volcanic rocks of Vesuvius; *2*, volcanic rocks of Somma; *3*, Tertiary and Quaternary clastic sediment; *4*, Mesozoic limestone; *5*, chilled margin and contact metamorphic halo of the magma chamber; *6*, differentiated magma of the AD 79 magma chamber that produced the white (w) and grey (g) pumice of the Plinian eruption.

is interrupted by three prominent surge beds. The progressive strength of the surge blasts is indicated by the increase in cast diameter of destroyed trees from 10 cm for the lower surge to 20 cm for the upper one. Above the grey pumice is a prominent sandwave horizon that contains a broken tree trunk 25 cm in diameter, and a mould of a tree 40 cm in diameter. Above the surge sequence are two pyroclastic flows. The lower one is rich in grey pumice and lacks stratification, the upper one has numerous reversely graded laminations, probably indicating that it is at its distal reaches[12]. Above the pyroclastic flows is a series of wet surge deposits with planar and massive beds rich in pisolites.

The Herculaneum section is ~20 m thick (Fig. 2). There is no pumice-fall horizon: at the base are sandwave surge beds. Above is a pumice-rich pyroclastic flow containing fragments of houses, carbonized wood and fumarolic pipes towards the top.

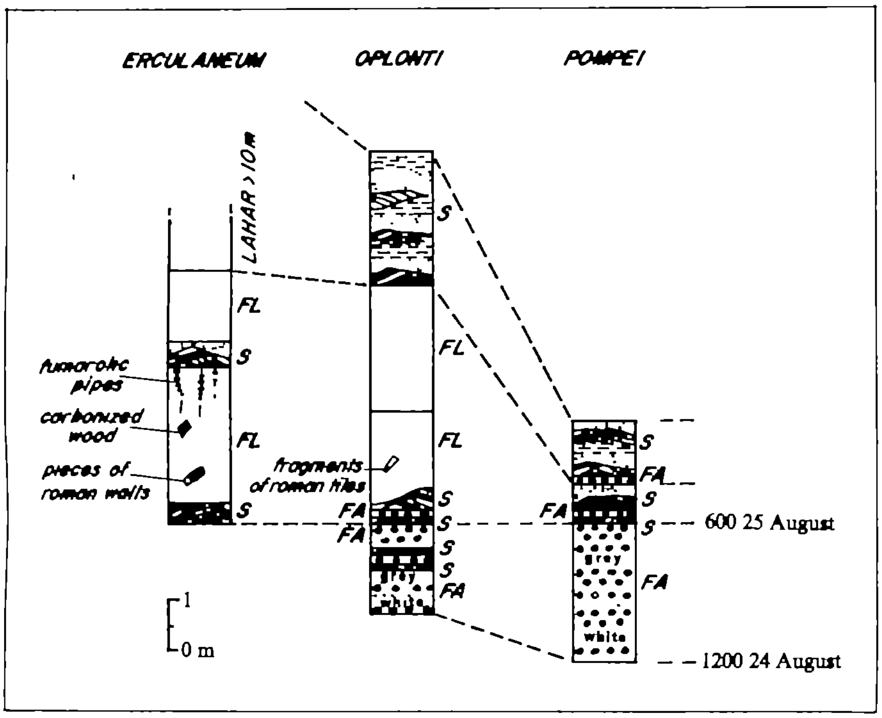


Fig. 2 Representative sections of AD 79 pyroclastic deposit from archaeological excavations. *FA*, air fall pumice; *S*, surge deposits, *FL*, pyroclastic flows.

The overlying sequence of sandwave surge deposits is followed by another pyroclastic flow. The remainder of the section is composed of lahars.

These deposits can be related to eruption phenomenology. A large volume (~2 km^3) of tephritic magma intruded a shallow (3–5 km deep) magma chamber within the Mesozoic limestone series[6–8]. The residence time of the magma was sufficient for a reaction skarn to form and a contact metamorphic halo that extended outwards into the carbonate country rocks to develop[7]. The metamorphic halo effectively sealed the differentiating magma system, although p_{CO_2} may have been significantly increased within the chamber as a result of decarbonation reactions in the carbonate wall rocks. A chilled magma margin may have originally formed, but with slow cooling and crystallization a coarser-grained syenite layer grew inwards from the upper part of the chamber as mafic cumulates settled to the bottom. The long period of differentiation produced a stagnant compositionally zoned upper zone of evolved magma rich in volatiles. The basic lower part was more homogenous, presumably due to circulation.

The eruption began (Fig. 3*a*) either when the gas pressure within the chamber became sufficient to rupture a conduit to the surface or, less likely, with the introduction of a new batch of hot magma from depth[13]. The initial eruption followed the model of Wilson[14–16] with decompression due to exsolving magmatic gas driving the gas thrust during the opening and widening of the conduit. Buoyant convective uplift of particles above the continuous gas thrust zone produced a persistent eruptive column from which the Plinian pumice-fall deposits were derived through lateral wind shear[17]. Using the average wind velocity at 9–11 km elevation (20 m s^{-1}) and terminal fall velocity data[5], a column height of ~17 km is computed. This corresponds well with the height of 16–26 km calculated from the theoretical formulae of Wilson *et al.*[16], assuming an average volume eruption rate of 4×10^4 m^3 s^{-1} for an 18-h period and an *F* value of 0.3–1.0.

The initial eruption produced the white highly fractionated grey pumice-fall with ~10 wt% lithic clasts in the approximate ratio of the stratigraphy traversed by the conduit. The eruption proceeded uninterrupted into the less-fractionated grey pumice-fall level with an increase in lithics and fraction of carbonates relative to other lithics and the appearance of felsic cumulates. This represents the removal of solidified magma and country rock surrounding the upper part of the magma chamber.

The greater density, larger size and wider dispersal of the grey pumice indicate a greater column height than for the white pumice. Lirer *et al.*[5] attribute this to tapping of a long (5 km) narrow (340 m diameter) magma chamber to provide sufficient increase in gas pressure to offset the decrease in gas content. The present data which indicate a magma chamber of ~2 km height and 1 km diameter, cast doubt on the above mechanism. Rather, the increased explosivity which widens the vent conduit system above the chamber could result from vaporization of a small mass ratio of external water to magma in the chamber during the grey pumice eruption. Minor surge layers within the upper part of the grey pumice deposit at Oplonti and other locations and an increase in abundance of lithic inclusions in the grey pumice support this interpretation.

The initial mass fraction of xenoliths in the white pumice was small and their ratio was proportional to their relative abundance in the stratigraphic column above the chamber. As the chamber became partially emptied during the eruption of the grey pumice, cavitation of the roof and the walls took place, (Fig. 3*b*) as indicated by the sharp increase in the proportion of deep inclusion types such as high-grade contact metamorphics, skarns and cognate plutonic nodules. The eruption may have become discontinuous at this stage.

Rupturing of the tight metamorphic encasement allowed interaction with the regional hydrologic system (Fig. 3*c*). Because the magma level within the emptying chamber was below the principal aquifer (Mesozoic carbonate sequence) an abundant source of water was available to the chamber when the venting pressure dropped below that of the hydrostatic head (~300–500 bar). The immediate effect was to generate strong hydromagmatic or phreatic explosions, depending on the degree of mixing of water and magma before decompression of the steam. The products included surges of fine-grained ashes, widespread phreatoplinian layers, vesiculated tuffs[18] and phreatic breccias.

At this stage, the production of small pyroclastic flows may have been directly related to contemporaneous hydromagmatic and magmatic activity occurring in different parts of the chamber. Marginal hydromagmatic explosions would have provided the fine fragmentation typical of pyroclastic flows[19–21] that is difficult to explain by other processes such as column collapse. Magmatic eruptions from the centre of the chamber would have produced the coarse pumice and provided a strong gas blast to maintain the eruption column. Homogenization within the chamber, vent and eruption column accounts for the rounding of the pumice and uniform texture characteristic of the pumice flows. Cooling of the system by water vaporization and enlargement of the vent due to strong explosions would have favoured column collapse.

At first hydromagmatic explosions alternated with purely magmatic activity; later they dominated. Because the eruptive system was gradually cooled through water vaporization during

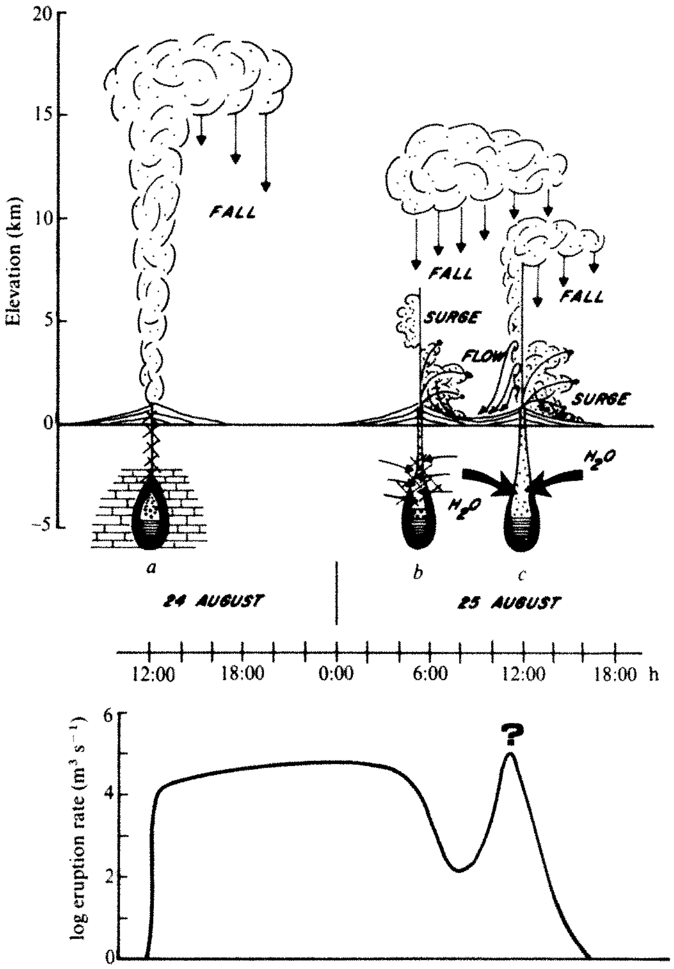


Fig. 3 Model of the AD 79 eruption of Vesuvius related to eruption phenomenology described by Pliney the Younger[23]. *a*, Gas pressure in the chamber ruptures a path to the surface producing an eruption column of about 17 km height. The eruptive rate during this phase averages 4.0×10^4 m^3 s^{-1}. *b*, A strong reduction in activity associated with cavitation of the chamber roof produces intermittent magmatic and hydromagmatic explosions. *c*, The terminal phase is dominantly hydromagmatic with pyroclastic flows and surges as the products. The eruption rate may reach 10^5 m^3 s^{-1}, close to the maximum theoretical rate for a prolonged maintained column[20].

the hydromagmatic phase, the products proceed through early hot dry surges and pyroclastic flows which are followed by colder wet surges and lahars and finally by phreatic explosion breccias. If the chamber had become completely evacuated of magma before significant hydromagmatic cooling, the final eruption would have been purely phreatic. These later products included xenoliths of syenite and mafic cumulates of biotite, olivine and pyroxene torn from the walls and floor of the chamber.

A similar pattern has been observed in four other Plinian eruption cycles identified at Vesuvius[22]: Avellino (3,500 yr BP), Gemelle (8,000 yr BP), Verdoline (13,000 yr BP) and Pomici di Base (17,000 yr BP). All began with a highly fractionated pumice-fall deposit in which the percentage of lithics increases with time. Those eruptions characterized by strong hydromagmatic activity (pyroclastic flow and surge) had a strong concentration of carbonate lithics in the pumice-fall immediately below the surge beds. This indicates an enlargement of the chamber roof within the regional aquifer by removal of the surrounding country rock. The surge deposits proceeded through an early hot dry type associated with pyroclastic flows to a final cold wet or even phreatic type.

The late-stage behaviour of the eruption depends both on internal magma dynamics, and on the hydrological framework. For example, the Verdoline magma chamber was located in the relatively impervious Tertiary flysch marls above the Mesozoic carbonate basement. Although there was an increase in amount of lithic clasts with time in the pumice fall deposit, the hydromagmatic phase was only minor. The other Plinian eruptions with magma chambers entirely within the carbonate basement, to the contrary, have produced extensive hydromagmatic and phreatic deposits.

This work was supported by USA–Italy cooperative research program of NSF, grant INT-7823984, and by the Italian Geodynamics Project, publication 319.

Received 6 May; accepted 31 October 1980.

1. Walker, G. P. L. & Croasdale, R. *J. geol. Soc. Lond.* **127,** 17–55 (1971).
2. Walker, G. P. L. *Geol. Rdsch* **62,** 431–446 (1973).
3. Sparks, R. S. J. & Wilson, L. *J. geol. Soc. Lond.* **132,** 441–451 (1976).
4. Nairn, I. A. & Self, S. *J. Volcan. geotherm. Res.* **3,** 39–60 (1978).
5. Lirer, L., Pescatore, T., Booth, B. & Walker, G. P. L. *Bull. geol. Soc. Am.* **84,** 759–772 (1973).
6. Barberi, F. *et al. IUGG, XVII General Assembly,* Canberra (1979).
7. Barberi, F. & Leoni, L. *Bull. volcan.* (in the press).
8. Barberi, F. *et al. EEG 2nd Seminar on Geothermal Energy,* Strasbourg (1980).
9. Sheridan, M. F. & Updike, R. G. *Bull. geol. Soc. Am.* **86,** 571–581 (1975).
10. Wohletz, K. H. & Sheridan, M. F. *Geol. Soc. Am. Spec. Pap.* **180,** 177–193 (1979).
11. Self, S. & Sparks, R. S. J. *Bull. volcan.* **41,** 196–212 (1978).
12. Sheridan, M. F. *Geol. Soc. Am. Spec. Pap.* **180,** 125–136 (1979).
13. Sparks, R. S. J., Sigurdsson, H. & Wilson, L. *Nature* **267,** 315–318 (1977).
14. Wilson, L. *Geophys. J. R. astr. Soc.* **30,** 381–392 (1972).
15. Wilson, L. *Geophys. J. R. astr. Soc.* **45,** 545–556 (1976).
16. Wilson, L., Sparks, R. S. J., Huang, T. C. & Watkins, N. D. *J. geophys. Res.* **83,** 1829–1836 (1978).
17. Walker, G. P. L., Wilson, L. & Boswell, E. L. G. *Geophys. J. R. astr. Soc.* **22,** 377–383 (1971).
18. Lorenz, V. *Sedimentology* **21,** 273–291 (1974).
19. Walker, G. P. L. *J. Geol.* **79,** 696–714 (1971).
20. Sheridan, M. F. *J. geophys. Res.* **76,** 5627–5634 (1971).
21. Sparks, R. S. J. *Sedimentology* **23,** 147–188 (1976).
22. Delibrias, G., Di Paola, G. M., Rosi, M. & Santacroce, R. *R. Soc. Ital. Miner. Petrol.* **35,** 411–438 (1979).
23. Radice, B. *Pliny Letters and Panegyricus I* (Harvard University Press, Massachusetts, 1972).

Middle Ordovician chondrite in fossiliferous limestone from Brunflo, central Sweden

Per Thorslund

Department of Palaeontology, Uppsala University, Box 558,
S-751 22 Uppsala, Sweden

Frans E. Wickman

Department of Geology, University of Stockholm, Box 6801,
S-113 86 Stockholm, Sweden

The few discoveries of fossil meteorites have been attributed[1] to the ignorance about the appearance of meteorites among palaeontologists and other geologists who examine limestone quarries, coal mines, and other excavations in sedimentary rocks. We now report the first find of a fossil stony meteorite in Ordovician limestone. The meteorite is a chondrite, possibly an H-chondrite, and its terrestrial age is ~463 Myr. The fine structural details of the chondrules are often extremely well preserved, but the chemical composition and the mineralogy have changed dramatically, chromite being the only primary mineral preserved. The present major minerals are calcite, barite, a Cr–V– 'phengite' and a cobaltite-group mineral.

An earlier find reported by Yudin[2] consists only of isolated barred chrondrules in a Mesozoic bauxite. In 1952, a polished slab of Ordovician limestone was sent to one of us (P.T.) for examination because it contained an almost black clast about 10 cm across (Fig. 1). Thin sections were prepared of the foreign body and were examined by a petrographer, who reported that the clast was a metamorphic ultramafic rock consisting mainly of clinozoisite, calcite, magnetite and serpentine. This unusual discovery was explained by algae transporting the stone to its present position. However, as no similar ultramafic rocks were known on the continental (eastern) side of the Brunflo area the slab was set aside. The Siljan Ring, over 200 km to the south of Brunflo was one of Sweden's best areas for Ordovician and Silurian rocks, is an impact structure. In December 1979 the slab

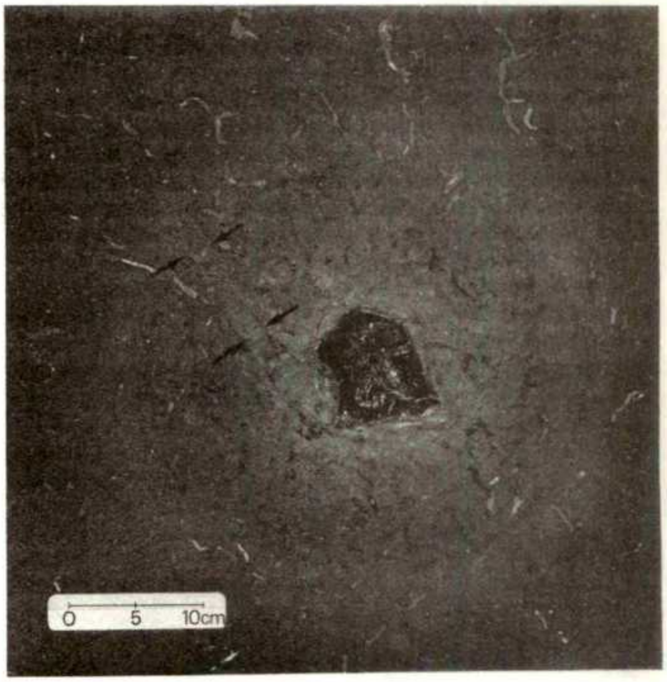


Fig. 1 The polished slab of limestone with the chondrite. The size of the slab is ~65 × 65 cm, and the meteorite ~10 cm across. The dark portions contain the structurally best preserved chondrules. The white portions are now calcite and barite. The cephalopod (arrowed) seems to have been struck by the meteorite. The differently coloured zones around the meteorite are visible.

Table 1 Electron microprobe analyses of chondritic chromites

	1	2	3	4	5	6	7	8	9
Cr_2O_3	56.8	56.9	56.1	54.4	55.9	60.8	57.2	56.7	57.2
Al_2O_3	6.5	5.9	5.3	5.7	6.0	3.9	6.3	5.9	5.8
V_2O_3	0.74	0.68	0.72	0.73	0.65	0.70	0.71	0.68	0.69
TiO_2	1.73	2.33	2.81	3.23	2.72	1.47	2.12	2.27	2.47
FeO	30.9	31.2	33.0	34.5	30.1	30.5	30.6	31.3	30.8
MgO	1.98	2.66	1.99	1.62	2.72	2.32	2.73	2.51	3.0
MnO	0.50	0.94	0.74	0.63	1.03	0.79	0.86	0.99	0.81
ZnO	0.93	—	—	—	1.09	—	—	—	—
Total	100	100.61	100.66	100.81	100.21	100.48	100.52	100.35	100.77

Analyses 2–9 from ref. 4. Number in parenthesis gives number of meteorites analysed for 2–4, and number of grains analysed for 1, 5 and 6–9.
Analyses: 1, average present find, normalized to a sum of 100 after exclusion of 0.33 wt% CaO (11); 2, average H5 and H6 (10); 3, average L5 and L6 (7); 4, average LL5 and LL6 (6); 5, average Pultusk (4); 6, average H3 (12); 7, average H4 (21); 8, average H5 (28); 9, average H6 (11).

with its clast was re-examined with the idea that it could be a stony meteorite; the presence of chrondrule-like boundries was noticed.

The meteorite occurs in a reddish brown limestone from the Rödbrottet quarry (lat 63°7′ N, long 14°17′ E), near Brunflo, not far from Östersund, central Sweden. Determination of the conodonts in the slab shows that the meteorite fell during the Aseri Stage of the Middle Ordovician (~463 Myr ago).

In a ~3 cm-wide border zone, delicate structural details of chondrules are preserved, although the primary minerals have been replaced. Two examples are shown in Fig. 2, one barred chondrule and one radiating chondrule. The chondrules are well defined in a matrix, giving the impression of the petrologic types 3–5 of Van Schmus and Wood[3].

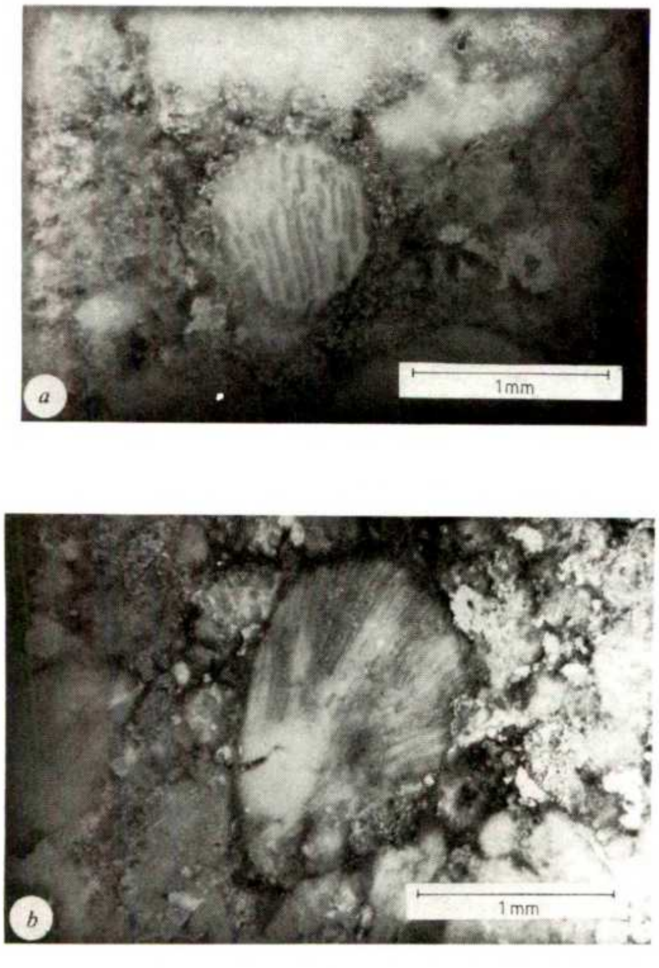


Fig. 2 Examples of well preserved chrondrules from the polished surface of the slab: *a*, a barred chondrule; *b*, a radiating chondrule.

The present mineralogy of the chondrite is quite different from the original one (and from that given by the petrographer). Olivine, pyroxene, nickel–iron and troilite no longer exist. In fact iron and magnesium have been lost almost completely from the meteorite. Instead, the dominating mineral is calcite, the other main constituents being barite, a Cr–V-bearing 'phengite' and a cobaltite—group mineral with more cobalt than nickel. Electron microprobe experiments reveal that only one primary mineral seems to have been preserved, chromite.

Table 1 shows some chromite analyses. The present chromite is seen to be a typical chrondrite chromite, with a homogeneous composition. The calcium content is interpreted as belonging to the calcite. The present chromite contains 0.93 wt% ZnO; the only chondrite found by Bunch *et al.*[4] to contain zinc-bearing chromite is the H5 Pultusk. It is very difficult to make a definite classification of the find, but it seems to be most similar to the H group, most likely the subgroups H4–H5. However, note that analysis 1 in Table 1 could be interpreted to indicate an unknown group of chrondrites.

Around the meteorite are concentric rings in different shades of brown, red, grey and green. They clearly represent reactions between the carbonate mud or, if later, limestone and the original meteoritic matter. These zones have not as yet been studied.

A badly preserved orthocone nautiloid cephalopod, which occurs in contact with the meteorite, could have been hit and killed by the meteorite.

Much of the discussion of the sources of chrondritic meteorites focuses on the part played by Apollo and Amor objects in their relationships to comets. Wetherill's results[5] indicate that for bodies in Earth-crossing orbits one half is eliminated in the first 10 Myr but that the rate of elimination then slows down greatly. He has shown that in steady state conditions and a population of ~800 Apollos (15 added per Myr) their median age will be ~200 Myr and ~35% of them will be more than 500 Myr old. Even if these figures were considered with great caution the present find is not inconsistent with Wetherill's conclusions.

The concentration of cosmic ray exposure ages around 4 Myr for modern H chondrites—this indicates that they come from another source than did the clast we have studied. Of course, it has been assumed that the composition of the chromite corresponds to those of H chrondrites and not to some unknown group. A detailed description of this first definite fossil stony meteorite will be given elsewhere[6].

Received 29 September; accepted 2 December 1980.

1. Nininger, H. H. *Find a falling star* (Eriksson, New York, 1972).
2. Yudin, I. A. *Meteoritics* **6,** 99–103 (1971).
3. Van Schmus, W. R. & Wood, J. A. *Geochim. cosmochim. Acta* **31,** 747–765 (1967).
4. Bunch, T. E., Keil, K. & Snetsinger, K. G. *Geochim. cosmochim. Acta* **31,** 1569–1582 (1967).
5. Wetherill, G. W. *Yb. Carnegie Instn. Wash.* **77,** 456–471 (1978).
6. Thorslund, P. & Wickman, F. E. *Lithos* (in preparation).

Direct use of dissolved organic carbon by agglutinated benthic foraminifera

T. E. DeLaca*‡, D. M. Karl† & J. H. Lipps*

* Department of Geology, University of California, Davis, California 95616
† Department of Oceanography, University of Hawaii at Manoa, 2525 Correa Road, Honolulu, Hawaii 96822

Foraminifera are known to obtain nutrients in a variety of ways: they are omnivores, carnivores or herbivores[1-3] and some species are known to use the extracellular metabolites of their photosynthetic endosymbionts[1,2,4]. None, however, has previously been proven to utilize exogenous dissolved organic carbon directly, although this is well known in other marine species[5,6]. Knowledge of foraminiferal trophic positions is important because foraminifera are common in most marine communities[1] and may be the most abundant eukaryotic organism in the extensive deep-sea benthos[7-9]. Our studies of benthic foraminifera from an unusual Antarctic shallow water embayment now show that certain species utilize both particulate and dissolved organic material in their nutrition.

Our studies were conducted at McMurdo Sound, Antarctica, located in the southwest corner of the Ross Sea at latitude 77 °S and longitude 166°E (Fig. 1). The Sound is approximately 50 km wide; however, its eastern and western sides differ greatly in water column productivity and benthic biomass due to local oceanographic conditions[10] (see Fig. 1). The eastern side of McMurdo Sound is extremely productive during the austral summer, whereas the western side at New Harbor has very low productivity. The western side is rich in nitrate, phosphate and silica and low in plankton and particulate organic material (ref. 11 and D.M.K., unpublished data). Based on water analyses, species diversity and faunal affinities, New Harbor is more similar to the deep sea than to other shallow-water areas around Antarctica[11]. The organisms we discuss here were found only at New Harbor. Our investigation covered an area of 15–40 m depth, with a temperature of −1.8 °C. The sediments were rich in organic material. Analyses of dissolved amino acids, as shown by fluorometric analyses[12] of water collected in dialysis bags, demonstrated maximum concentrations in the sediment (20–30 μM at 1 cm) and smaller concentrations in the water column (5 μM at the sediment water interface, 1 μM at 10 cm from the bottom). This trend in the distribution of amino acids is similar to that of total dissolved organic carbon in the deep sea[13].

Our studies focused on *Notodendrodes antarctikos*[14] and an as yet undescribed species. Both are arborescent agglutinated foraminifera in the superfamily Ammodiscacea (Fig. 2). These large foraminifera are conspicuous components of the bottom assemblages at New Harbor, with abundances of >200 individuals per m^2 (ref. 14). Our initial studies on the physiology and trophic positions of these organisms revealed that they use captured suspended particulate organic material and dissolved organic materials in their nutrition. As seen by scanning electron microscopy of specimens which were fixed *in situ* and later dried using a critical-point technique, suspended particles are trapped in sticky cytoplasmic masses which extend from the branches of the tests. The foreign material caught in pseudopodia consists of benthic diatoms, bacteria, amorphous flocculent organic and inorganic material (such as silt, sand, broken diatom frustules and sponge spicules) with adhering bacteria. This material is probably suspended by the activities of larger invertebrates. Marine bacteria, labelled with ^{3}H-adenine (in the laboratory) were presented to foraminifera *in situ*, and these foraminifera were later fixed *in situ*. Microautoradiography revealed that the labelled bacteria are captured, digested and incorporated into the protoplasm of the organisms. However, the capture of particulate matter, in natural conditions, seems to occur infrequently in the population studied (3 of 32 specimens).

The foraminifera were tested for their ability to take up dissolved organic carbon by exposing them to several radiolabelled substrates, including amino acids (algal protein hydrolysate, Amersham), glucose and adenine (Table 1). We found that they took up the first two but did not absorb adenine. As bacterial contamination in our experiments could have produced erroneous results, we used ^{3}H-adenine (rapidly taken up by most bacteria[13]) in some of our experiments as a control in dual-label experiments with ^{14}C-labelled substrates. When tritium was detected in liquid scintillation analyses, those samples were assumed to be contaminated and were eliminated. 'Time zero' and NaCN- killed controls were used in all experiments.

Experiments measuring the uptake of dissolved organic carbon were conducted in the laboratory. Foraminifera cleaned in filter-sterilized sea water (0.22 μm) were placed in 10 ml of filtered (0.22 μm) sea water (1.0 μm free amino acids in solution) containing 2.5 μCi of ^{14}C-UL-protein hydrolysate

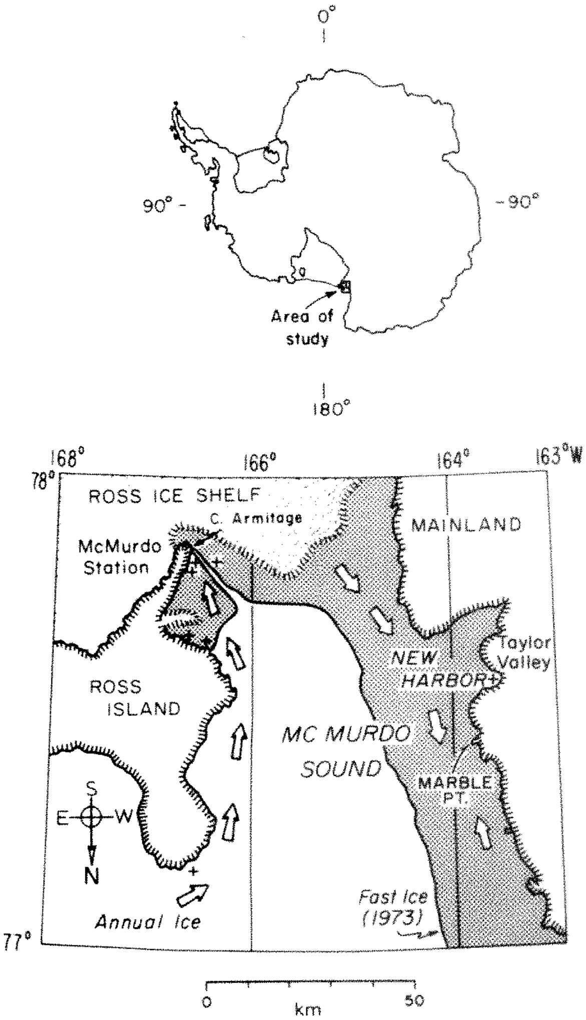


Fig. 1 Antarctica, showing the location of the study area. [During past investigations in McMurdo Sound we have surveyed several localities (+).] The two new species of foraminifera discussed in the present report were found only at New Harbor. Arrows indicate the probable direction of water movement. In the eastern and central parts of the Sound currents move south, advecting large standing crops of plankton from more northerly waters. The annual break up of sea ice on the eastern side allows local productivity to increase dramatically. In contrast to the highly productive eastern side, the western side of the Sound is bathed by northward moving currents[12] which probably originate under the Ross Ice Shelf and are poor in organic nutrients.

---

‡ Present address: Marine Biology Research Division A-002, Scripps Institution of Oceanography, La Jolla, California 92093.

0028-0836/81/040287—03$01.00

(0.0165 μmol free amino acids). The final concentration of free amino acids in solution was 2.65 μM, and incubation times of 0.58, 1.08, 2.25, 3.25 and 4.17 h were used, all experiments being conducted at −1.0 °C. Following incubation, the foraminifera were washed and: (1) solubilized in hot and cold 10% trichloroacetic acid (TCA), (2) placed in sealed reaction flasks with 10 ml of filtered sea water for measurement of $^{14}CO_2$ evolution, or (3) fixed for autoradiography.

Cellular contents of solubilized foraminifera were analysed using liquid scintillation. Uptake rates of 6.14×10^{-4} μmol h^{-1} (s.d. $= 4.26 \times 10^{-4}$, $n = 20$) were found. The tests of individual organisms were subsequently removed, washed three times in hot and cold 10% TCA to remove any remaining soluble material and re-analysed. Up to 41.3% of the label was incorporated into non-soluble material that remained in the test.

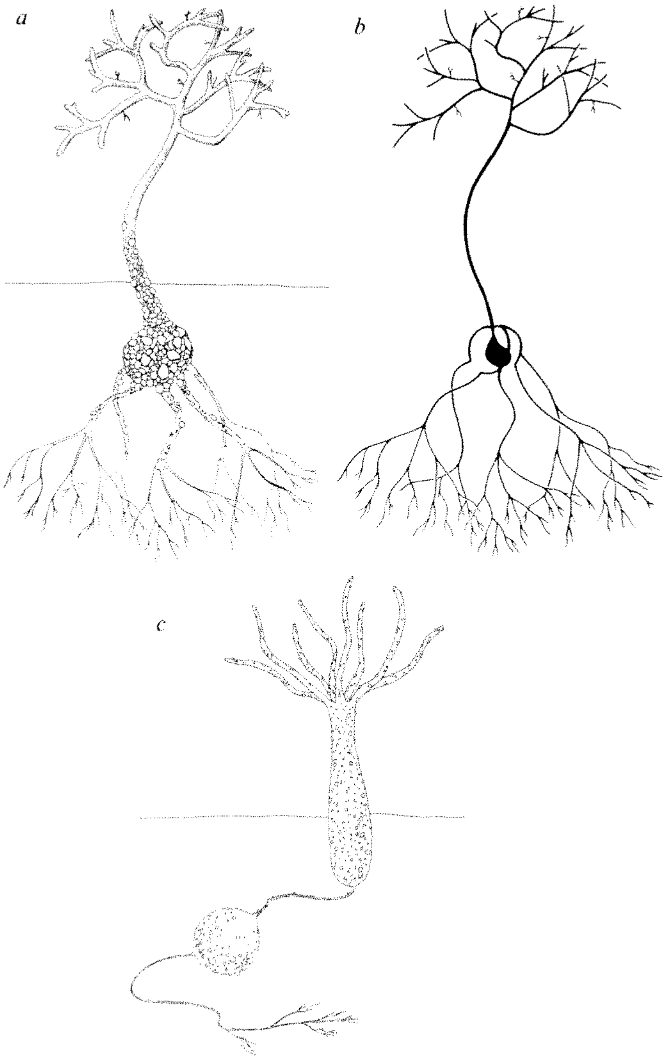


Fig. 2 *a*, *Notodendrodes antarctikos*[14] drawn in life position. The branching upper half extends above the sediment-water interface and the bulb and root system are buried in the sediment (average length 20 mm). *b*, The protoplasm of this species as would be seen if the test was removed. Cytological studies have shown that approximately one-third of the protoplasm (containing the nucleus) resides within the bulb. About one-third of the cytoplasm extends up the stem and occupies the branching system. Small pseudopods extend from the branches and upper portions of the stem. Particulate organic material is entrapped in these web-like projections. The remaining one-third of the cytoplasm resides in the extensive root system. Autoradiographic studies have shown that although dissolved organic material is taken up on all exposed cytoplasmic surfaces, the highest uptake occurs in the root system where the surface area is greatest. *c*, An unnamed agglutinated foraminfer (Notodendrodidae gen. et sp. n.) drawn in life position (average size 18 mm).

Table 1 Summary of uptake

Protein hydrolysate	
Total uptake	6.14 μmol h^{-1}
Incorporation into TCA-insoluble form	41.3% h^{-1}.
$^{14}CO_2$ liberated in 12 h	10%
Glucose	
Total uptake ^{14}C-glucose	6.1×10^{-2} μmol h^{-1}
$^{14}CO_2$ liberated in 12 h	46%
Adenine	
Total uptake	Undetectable

Experiments were conducted at −1 °C. Final concentration of free amino acid in laboratory experiments was 2.65 μM. Concentration of free amino acid in the sediment where the organisms are found was found to be 20–30 μM. See text for further explanation. Concentrations of glucose in the environment and laboratory experiments were not adequately measured.

The labelled foraminifera placed in reaction vials were incubated for 12 h at −1°C. The incubation was terminated with the addition of 0.2 ml of ethanolamine (placed on filter paper in the reaction vessel) and 1 ml H_2SO_4 (0.5 M) was injected into the seawater. After 2 h at room temperature (23 °C), the ethanolamine and filter paper were analysed by liquid scintillation. The results demonstrated that 10.0% of the total ^{14}C taken up was respired.

The results of autoradiography showed that the foraminifera took up dissolved organic carbon through any cell surface; consequently, uptake was a function of surface area. The 'root system', with its large surface area, was the principal site of uptake in the organisms analysed.

These two species of foraminifera seem to have ecological and physiological adaptations which are well suited to a highly seasonal oligotrophic environment. At New Harbor, primary productivity is very low and limited to a period of 1 or 2 months during the year, during which time the growth of a diatom flora on the undersurface of the 3–4-m thick sea ice or on the sediments of very shallow water (< 10 m) provides sparse organic input. The under-ice flora rains down to the bottom, where a portion is directly utilized while the remainder enters the detritus food web. Decomposition of this sedimented organic material generates relatively high concentrations of dissolved organic carbon in the sediment and supports a large and metabolically active microbial population (D. M. K., unpublished data). These foraminifera can capture and digest this particulate material either as it settles to the bottom or as it is resuspended by deposit-feeding invertebrates, although capture of particles seems to occur only infrequently. Dissolved organic carbon may contribute significantly to the nutrition of these foraminifera. During periods of no primary productivity, ~10 months of the year, they may sustain themselves on the high dissolved organic carbon reserves which are continually generated by the activities of heterotrophic microorganisms in the sediment.

Our results demonstrate that foraminifera are able to use both saprozoic and holozoic (phagotrophic) modes of nutrition. The amount and type of organic material available to foraminifera are therefore both much greater and more diverse than was previously thought. Re-evaluation of current ideas of the trophic positions of these extremely abundant organisms in benthic ecosystems is necessary, especially in deep-sea communities where they have been found to represent the highest test-free biomass of all eukaryotic organisms[9].

Present studies are analysing the metabolic rates of these organisms and will determine the proportion of their total energy requirements supplied by dissolved organic carbon.

We thank Professor A. E. V. Haschemeyer for her assistance in experimental design and encouragement, Drs W. L. Stockton and R. R. Hessler for reading the manuscript and Ms E. Bailey for drawing the figures. This research was funded by NSF grant DPP 76-17231.

Received 14 April; accepted 24 October 1980.

1. Boltovsky, E. & Wright, R. *Recent Foraminifera* (The Hague, 1976).
2. Lee, J. J. in *Foraminifera* (eds Hedley, R. H. & Adams, C. G.) 207 (Academic, London, 1974).
3. Lee, J. J. *et al. J. Protozool.* **13,** 659–670 (1966).
4. Smith, D. F. & Wiebe, W. J. *Aust. J. Fish. Fresh. Res.* **28,** 311–319 (1977).
5. Jorgensen, C. B. *Biol. Rev.* **51,** 291–328 (1976).
6. Stevens, G. C. in *Symp. Nitrogen Metabolism and the Environment* (eds Campbell, J. W. & Goldstein, C.) 155–184 (Academic, New York, 1972).
7. Hessler, R. R. in *The Biology of the Oceanic Pacific* (ed. Miller, C. B.) 79–93 (Oregon State University Press, Corvalis, 1974).
8. Tendal, O. & Hessler, R. R. *Galathea Rep.* **14,** 165–194 (1977).
9. Smith, K. L. Jr *Mar. Biol.* **47,** 337–347 (1978).
10. Littlepage, J. L. *Antarct. Res. Ser.* **14,** 1–37 (1971).
11. Dayton, P. K. & Oliver, J. S. *Science* **197,** 55–58 (1977).
12. North, B. B. *Limnol. Oceanogr.* **20,** 20–27 (1975).
13. Karl, D. M. *Appl. envir. Microbiol.* **38,** 850–860 (1979).
14. DeLaca, T. E., Lipps, J. H. & Hessler, R. R. *J. Linn. Soc. Lond. Zool.* **69,** 205–224 (1980).

Lichen recolonization in London's cleaner air

C. I. Rose* & D. L. Hawksworth†

* Department of Applied Biology, Chelsea College, University of London, Hortensia Road, London SW10 0QX, UK
† Commonwealth Mycological Institute, Ferry Lane, Kew, Surrey TW9 3AF, UK

Mean sulphur dioxide levels have fallen markedly in London, UK during the past 15 yr. Although lichens growing on trees are particularly sensitive to this air pollutant, there are few studies of lichen recolonization following pollution episodes. A survey of 29 sites in the north and west of Greater London, reported here, demonstrates that several species extinct or very rare in the area in 1970 have extended their ranges considerably. Studies on growth rates of the species concerned suggest that many of the sites discovered have been recolonized within the past 3–7 yr. If current trends continue, further improvements in the lichen flora can be expected in the next few years. However, it is unlikely that London will regain in the forseeable future many of the species lost during the past two centuries.

In the British Isles, smoke and sulphur dioxide (SO_2) levels have fallen by ~80% and 50% respectively in urban areas since 1960[1]; in Inner London, the mean annual level of SO_2 declined from 200–250 $\mu g\,m^{-3}$ to $<130\,\mu g\,m^{-3}$ in the same period[2]. Recorded changes in mean annual and mean winter SO_2 values at six stations in north and west London are shown in Fig. 1. Lichens have been used extensively as indicators of air pollution and have proved to be of particular value in estimating levels of SO_2 pollution in temperate regions[3–8]. Most studies have been concerned with changes in conditions of increasing levels of air pollutants but there are a few scattered reports of improving lichen vegetation in ameliorating conditions from Germany[9], Sweden[10,11] and several urban areas of the British Isles—Birmingham[12], Leeds[13–15], Newcastle-upon-Tyne[16] and Sheffield[17]. Some increases in the abundance of the SO_2-tolerant *Lecanora conizaeoides* between 1953 and 1967 at Holland Park, London, have been reported[18] and there are a few recent records of other more sensitive species from London, including a single plant of *Hypogymnia physodes* at Brent Reservoir in 1972 (ref. 19), and *Parmelia sulcata* on a wall at Brook Green in 1976 by P. W. James (personal communication). No detailed survey of the recolonization of lichens on trees in a large urban and suburban area has, however, been previously documented in the British Isles.

To determine what, if any, improvements of the lichen flora of the London area had taken place, 29 sites in the north and west of Greater London were studied between December 1979 and February 1980. The number of trees available for study varied but a minimum of five was examined at each site. Lichen abundance was estimated on a 1–5 scale (Table 1) and particular attention was paid to *H. physodes*, *Evernia prunastri*, *Parmelia caperata*, *Parmelia subaurifera*, *P. sulcata* and *Usnea subfloridana*, as these species have different ranges of tolerance to SO_2 (ref. 3) and are easily identified even when only a few millimetres in size. Additional interesting species were encountered in some sites, especially the Ruislip Local Nature Reserve, where 25 lichens not recorded from there in 1963 were found, including nine which are probably recent colonists (these results are presented in detail elsewhere[20]). Other notable records include *Buellia punctata* and *Xanthoria polycarpa* on dead elms near Northwood Hills (national grid ref. TQ 104898), the former also

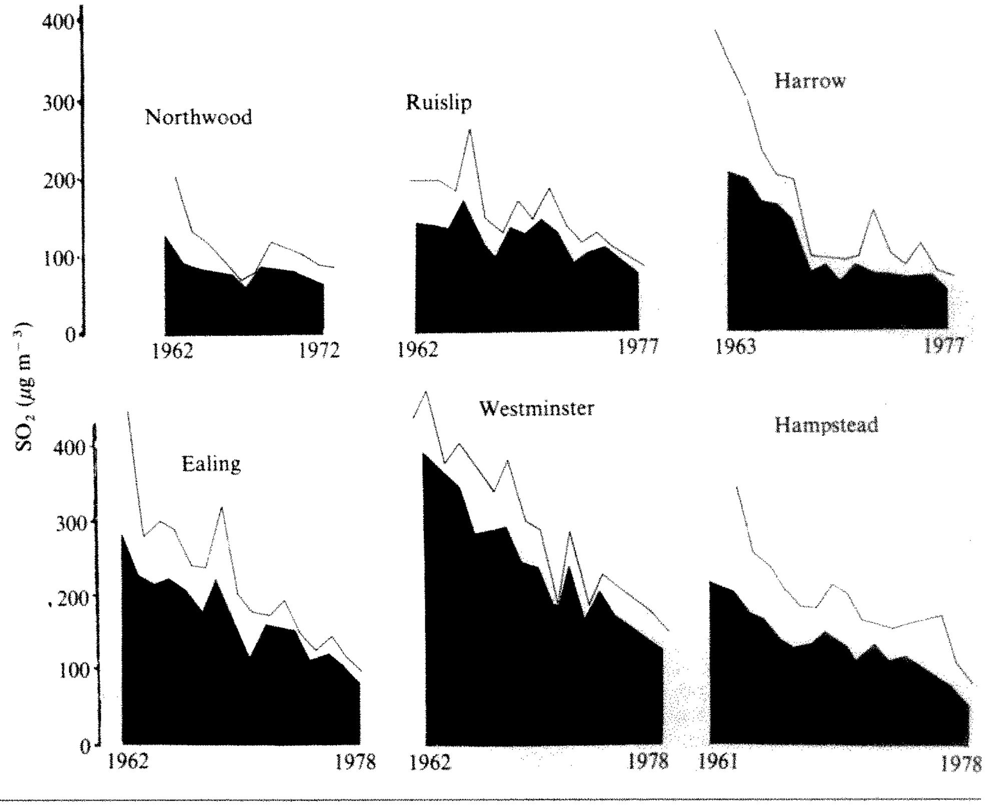


Fig. 1 Changes in mean annual (black) and mean winter (unshaded) SO_2 values at six recording stations (see Fig. 3) in north-west London. Based on data compiled by the Warren Spring Laboratory[36].

0028–0836/81/040289—04$01.00

Table 1 Abundance of six lichen species at 29 sites in north-west London in 1979–80 compared with their status in Inner London in 1970‡

Site	National grid ref.	*Evernia prunastri*	*Hypogymnia physodes*	*Parmelia caperata*	*Parmelia subaurifera*	*Parmelia sulcata*	*Usnea subfloridana*	Lichen zone*
Ruislip LNR†	TQ 090901	4	3	1	4	5	4	6
Ruislip Lido†	TQ 088906	3	3	—	2	4	2	5
Copse Wood†	TQ 084901	1	3	—	—	2	—	5
Park Wood†	TQ 093905	1	3	—	—	2	—	5
Brent Reservoir	TQ 217884	1	1	—	—	—	—	5
Hampstead Heath (Wildwood Road)	TQ 262886	—	2	—	—	1	—	5
Hampstead Heath (Vale of Health)	TQ 278874	1	1	—	—	2	—	5
Harrow Weald Common (south)†	TQ 139934	—	2	—	—	1	—	5
Mad Bess Wood†	TQ 075906	—	3	—	—	1	—	4
Bayhurst Wood†	TQ 068891	1	1	—	1	1	—	4
Pinner Park†	TQ 119895	—	1	—	—	—	—	4
Horsenden Hill	TQ 163854	—	1	—	—	—	—	4
Harrow Hill Pond	TQ 155870	1	1	—	1	1	—	4
Harrow Weald Common (north)†	TQ 143933	—	1	—	—	—	—	4
Ickenham Marsh†	TQ 088869	—	—	—	1	1	—	4
Hanwell Golf Course Pond	TQ 142813	—	1	—	1	1	—	4
Osterley Park	TQ 147795	—	2	—	2	2	—	4
River Thames at Kew	TQ 175777	—	—	—	—	—	—	3
River Thames at Isleworth	TQ 173777	—	—	—	—	—	—	3
River Brent at Perivale	TQ 170833	—	—	—	—	—	—	3
Lyttleton Fields	TQ 262891	—	—	—	—	—	—	3
Coldfall Wood	TQ 276918	—	—	—	—	—	—	3
Old Park Wood†	TQ 045926	—	—	—	—	—	—	3
West Brompton Cemetry	TQ 257778	—	—	—	—	—	—	3
River Crane at Twickenham†	TQ 133733	—	—	—	—	—	—	3
Holland Park	TQ 246802	—	—	—	—	—	—	3
Wormwood Scrubs	TQ 214823	—	—	—	—	—	—	3
Hyde Park	TQ 270800	—	—	—	—	—	—	2
Regents Park	TQ 280830	—	—	—	—	—	—	2
Status in London in 1970 (ref. 18)		Extinct; last seen ~1800	Rare or extinct; last seen 1957 at Mill Hill	Extinct; last seen ~1800	Extinct; last seen late 1800s	Rare; a single record on concrete in 1968 at Wood Green	Genus last seen ~1800	

Sites are arranged according to their rating on the Hawksworth and Rose[3] scale. Most species occurred on *Quercus* and *Salix*, but records from *Betula* and *Carpinus* are also included; nutrient-rich bark is not considered.

* Zone to which the richest part of the site is referred on the 0–10 scale of Hawksworth and Rose[3].

† Outside the 16-km-radius circle centred on Trafalgar Square used in the study of Laundon[18].

‡ Abundance estimated on a 1–5 scale: 1, cover under 1% (1–5 thalli per tree); 2, 1–5% (5–20 thalli); 3, 5–20% (20–100 thalli); 4, 20–50% (100–1,000 thalli); 5, 50–100% (over 1,000 thalli).

in Pinner Park (TQ 135906) on dead elm stumps and *Rinodina exigua* on poplar in Eastcote (TQ 107889).

Comparison of the data presented in Table 1 with that of the same species in London in 1970 (ref. 18) leaves no doubt that there has been a marked incursion into north-west London in the past decade. To estimate the dates at which recolonization might have occurred, growth rates were studied by making direct area measurements, tracing 30 thalli of each of *H. physodes*, *P. sulcata* and *P. subaurifera* between April 1979 and March 1980 in the Ruislip Local Nature Reserve. These studies (to be reported in detail elsewhere) showed that the increases in area per unit time were significantly different (t-test between $P = 0.5$ and $P = 0.001$) between these three species. The annual increases in area at this site were 100–150% for *H. physodes*, 215–230% for *P. subaurifera* and 160–200% for *P. sulcata*, but no correlation between starting size and percentage increase per unit time was obtained (*H. physodes*, $R = -0.072$; *P. subaurifera*, $R = 0.019$; *P. sulcata*, $R = -0.002$; at $P = 0.05$ not significantly different from zero). In general, relative growth of circular lichen thalli expressed on a percentage basis is expected to be constant during the exponential growth phase, but to decline thereafter[21]; the exponential phase tends to last 10–15 yr and the Ruislip data therefore conform to the expected pattern. Reports of the effect of air pollutants on the growth rate of lichens are contradictory and the effect may vary both with the species and pollutant levels involved. However, a study on *Parmelia saxatilis* (a species closely related to *P. sulcata*) in north-east England indicated that growth was not significantly affected by substantial increases in ambient SO_2 levels[22]. Furthermore, in species where the growth rate is depressed, the amount of depression involved would not materially affect the results reported here[23]. Variations in the growth rate of a species can also arise from differences in humidity or aspect, or whether specimens are growing on vertical as opposed to horizontal surfaces[21,24]. Based on the mid-value between the mean of

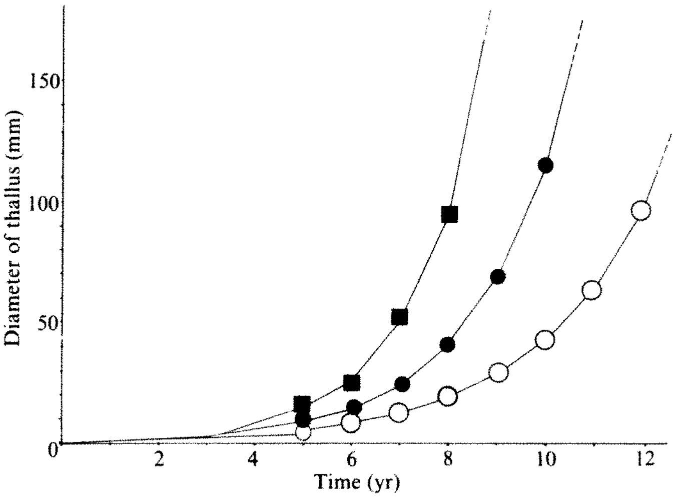


Fig. 2 Calibration curve used to estimate the minimum age of selected lichens from the diameter of their thalli. Based on the observed rates of increase in area per unit time at Ruislip Local Nature Reserve. ■, *P. subaurifera*; ●, *P. sulcata*; ○, *H. physodes*.

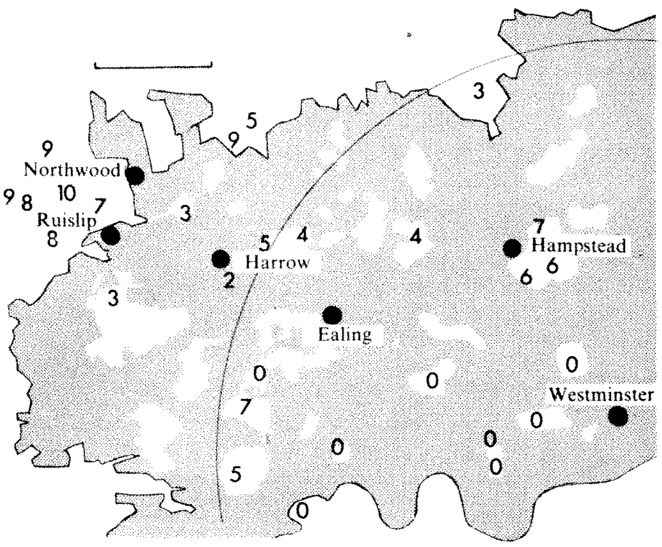


Fig. 3 North-west London showing the minimum age of foliose lichen colonies on trees (expressed in yr) calculated from observed thallus sizes and growth rates (Fig. 2). The arc indicates the limit of the survey by Laundon[18] (16-km radius from Trafalgar Square). Scale bar, 5 km.

readings for each species and the median of the range of values, the expected rate of increase in area per unit time was used to construct a calibration curve relating thallus diameter to age for these species (Fig. 2). This curve is designed to place a minimum age on individual colonies of the species concerned (where many specimens are present) rather than provide unequivocal dates for each individual thallus.

The results of the growth-rate study have been incorporated into Figs 3, 4, which show that recolonization by the foliose lichens considered has, subject to the above qualifications, taken place in the past 3–7 yr at most of the sites investigated, although this process was already in operation 8–10 yr ago in the Ruislip–Northwood and Harrow Weald areas. There is some collaborative data for these latter two areas: in Harrow Weald, populations of *H. physodes* declined between 1969 and 1971 but were still present at the latter date[25], while in the Ruislip Local Nature Reserve, foliose lichens were well developed in 1963 (ref. 20) (unfortunately there is no information on fluctuations in this site between 1963 and 1979). The results also show that the ranges of foliose lichens in the Ruislip area have extended considerably within the last 6 yr (Fig. 4).

As there is much evidence of the sensitivity of lichens to SO_2 pollution[3–8] and of the declines in SO_2 levels in the study area during the period when most of the lichen recolonization seems to have taken place, it is reasonable to suggest that the lichen recolonization now being seen is a direct response to falling mean SO_2 levels. Changes in levels of smoke would not be expected to cause responses in the lichen flora[26]. Air pollutants such as nitrogen oxides and ozone, which may have increased in the London area over the past two decades, would not have any adverse effects judging from experimental studies[27–29]. We therefore consider that these data demonstrate that London's lichens are responding to improvements in the air quality in the manner forecast by studies on deteriorating lichen floras. They also demonstrate that a cover of *L. conizaeoides* is insignificant in preventing the establishment of foliose lichens; indeed *H. physodes* and *P. sulcata* can establish directly on *L. conizaeoides*. Further, the lichenicolous fungus *Athelia arachnoidea*, which can kill many lichens and is especially common on *L. conizaeoides* in suburban areas, does not seem to oppose significantly re-establishment, as has been suggested[30]. It is also of interest that the sites in which recent recolonization has been found are not all contiguous. Recolonization has been more rapid in areas already with a nearby source of appropriate propagules, but dispersal has evidently been possible over considerable distances. Lichen diaspores are widespread in the air spora but soredia can also be carried on the feet of birds; as demonstrated at Brent Reservoirs[19], a site included in this survey.

The most suitable sites for recolonization are ones with well lit sheltered trees, in humid sites (for example, Ruislip Local Nature Reserve, Harrow Hill Pond, Wildwood Road), and often with willows overhanging a pond or marshy area surrounded by dense shrubs. In woodland sites, the earliest colonization tends to be at ground level, where air movements are less and the microclimate more favourable to growth; this would be expected as foliose lichens remain longest on tree bases in conditions of increasing SO_2 (ref. 3). In larger woodlands where the tree bases are very shaded and lichen growth is precluded, lichen communities including foliose species tend to develop on horizontal branches higher in the canopy where they can be easily missed. On willow, the first foliose lichen to colonize is *H. physodes*, followed by *P. subaurifera* and *P. sulcata*; *P. subaurifera* then tends to become more common (frequently growing where other species have died off) and *P. sulcata* dominant. On dry oak or birch, *H. physodes* seems better able to withstand competition from other species.

The investigations reported here concern only trees and studies in selected sites. We have not been able to undertake a tree-by-tree survey throughout north-west London and some recolonization sites will almost certainly have been overlooked. The sites listed in Table 1 are consequently representative rather than definitive as to the present position. Although the present survey was restricted to north-west London, the current improvement may be more general. For example, specimens of *Hypogymnia physodes* to 5 mm long were discovered on an oak stump near Pembroke Lodge in Richmond Park by F. S. Dobson in November 1980; this specimen was last recorded here in 1929 (ref. 17).

Lichens on man-made substrates, such as brick and asbestos-cement, can also be used as indicators of air pollutants but relatively little information is available on the way in which they respond, which is complicated by variations in the chemical nature of the substrate. However, our superficial observations suggest that they are responding in a parallel manner to the results presented here. In particular, *Lecanora muralis*, a species which changes its substrate range as SO_2 levels increase[23], can now grow on less basic substrates in the Kew area than it did in 1970; increases in this species have even led to comments in the local press arising from the general public's concern about this 'green fungus' on pavements[31,32].

[illegible] College
Upper Circular Road
Calcutta

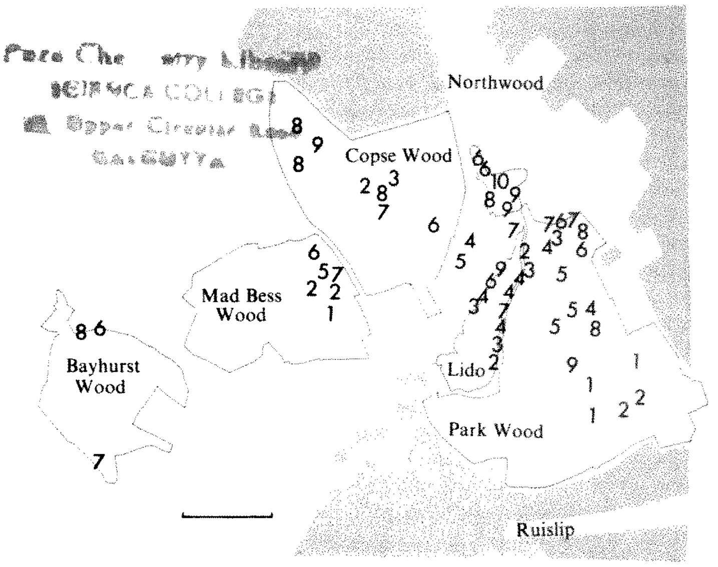


Fig. 4 Ruislip–Northwood Woods showing the minimum age of foliose lichen colonies on trees (expressed in yr) calculated from observed thallus sizes and growth rates (Fig. 2). Scale bar, 500 m.

The lichen flora of London has been studied since the early seventeenth century[33] and the historical records were combined into the survey of Laundon[18,34]. These records show a serious and progressive decline starting at least by the early nineteenth century, but which was particularly pronounced in the later decades of that century (for example, in Epping Forest[25]) and continued into the 1950s–60s. The loss of 129 lichen species from the area within a 16-km radius of Trafalgar Square has been attributed to increases in air pollution between 1800 and 1970 (ref. 18). Although the slight improvements reported here are encouraging, more drastic reductions in mean SO_2 levels are required before the environment is again suitable for many of the eliminated species. Indeed, a reduction in Inner London from the current mean of about 130 $\mu g\,m^{-3}$ to around 40–50 $\mu g\,m^{-3}$ would be necessary for survival of many of the more attractive foliose and fruticose species; transplant experiments have shown how hostile the Inner London environment is to some of these[35]. It has been estimated that mean SO_2 levels in urban areas will tend to fall until 1985, but thereafter return to 1975 levels due to changes in fuel-mix[1]. If this is true, only a limited rather than a substantial improvement in the lichen flora of the area seems likely.

We thank Mr P. W. James, Mr M. P. Jones, Mr J. R. Laundon and Drs C. J. Muskett and F. Rose for help. This work was carried out while C.I.R. was in receipt of a Chelsea College (University of London) Research Studentship.

Received 2 September; accepted 12 November 1980.

1. *The United Kingdom Environment 1979: Progress of Pollution Control,* Pollut. Pap. no. 16 (HMSO, London, 1979).
2. Sanford, H. *Lond. Nat.* **58,** 89–92 (1979).
3. Hawksworth, D. L. & Rose, F. *Nature* **227,** 145–148 (1970).
4. Ferry, B. W., Baddeley, M. S. & Hawksworth, D. L. (eds) *Air Pollution and Lichens* (Athlone Press of the University of London, London, 1973).
5. Hawksworth, D. L. *Lichenologist* **6,** 122–125 (1974); **7,** 62–66, 173–177 (1975); **8,** 87–91, 179–182 (1976); **9,** 77–82, 147–151 (1977); **10,** 95–100 (1978).
6. Hawksworth, D. L. & Henderson, A. *Lichenologist* **10,** 227–230 (1978); Henderson, A. & Hawksworth, D. L. *Lichenologist* **11,** 91–95 (1979).
7. Henderson, A. *Lichenologist* **11,** 313–319 (1979); **12,** 145–148, 397–402 (1980).
8. Hawksworth, D. L. & Rose, F. *Lichens as Pollution Monitors,* Stud. Biol. no. 66 (Arnold, London, 1976).
9. Jürging, P. *Biblthca lich., Lehre* **4,** 1–164 (1975).
10. Skye, E. & Hallberg, I. *Oikos* **20,** 547–552 (1969).
11. Westman, L. *Wahlenbergia* **2,** 1–146 (1975).
12. Lindsay, D. C. *Proc. Bgham nat. Hist. Soc.* **23,** 234–249 (1978).
13. Seaward, M. R. D. *Proc. Leeds lit. phil. Soc., Sci.* **10,** 141–208 (1976).
14. Henderson, A. *Naturalist, Hull* **1977,** 141–144 (1977).
15. Henderson-Sellars, A. & Seaward, M. R. D. *Environ. Pollut.* **19,** 207–213 (1979).
16. Gilbert, O. L. *Lichenologist* **12,** 325–395 (1980).
17. Hawksworth, D. L. *Lichenologist* **4,** 105–193 (1969).
18. Laundon, J. R. *Lond. Nat.* **49,** 20–69 (1970).
19. Bailey, R. H. & James, P. W. *Lichenologist* **11,** 105–106 (1979).
20. Hawksworth, D. L. & Rose, C. I. *J. Ruislip Distr. nat. Hist. Soc.* **22,** 23–29 (1979).
21. Topham, P. B. in *Lichen Ecology* (ed. Seaward, M. R. D.) 31–68 (Academic, London, 1977).
22. Gilbert, O. L. *Lichenologist* **5,** 11–17 (1971).
23. Seaward, M. R. D. in *Lichenology: Progress and Problems* (eds Brown, D. H., Hawksworth, D. L. & Bailey, R. H.) 323–357 (Academic, London, 1976).
24. Armstrong, R. A. thesis, Univ. Oxford (1974).
25. Hawksworth, D. L., Rose, F. & Coppins, B. J. in *Air Pollution and Lichens* (eds Ferry, B. W., Baddeley, M. S. & Hawksworth, D. L.) 330–367 (Athlone Press of the University of London, London, 1973).
26. Hawksworth, D. L. in *Air Pollution and Lichens* (eds Ferry, B. W., Baddeley, M. S. & Hawksworth, D. L.) 38–76 (Athlone Press of the University of London, London, 1973).
27. Brown, D. H. & Smirnoff, N. *Lichenologist* **10,** 91–94 (1978).
28. Nash, T. & Sigal, L. *Bryologist* **82,** 280–285 (1979).
29. Nash, T. *Bryologist* **79,** 103–106 (1976).
30. Arvidsson, L. *Svensk bot. Tidskr.* **72,** 285–292 (1979).
31. *Brentford & Chiswick Times,* 11 January (1980).
32. Gilbert, J. L. *Bull. Br. Lichen Soc.* **46,** 7 (1980).
33. Hawksworth, D. L. & Seaward, M. R. D. *Lichenology in the British Isles 1568–1975* (Richmond Publishing, Richmond, 1977).
34. Laundon, J. R. *Lichenologist* **3,** 277–327 (1967).
35. Ferry, B. W. & Coppins, B. J. *Lichenologist* **11,** 63–77 (1979).
36. Warren Spring Laboratory *The Investigation of Air Pollution. National Survey. Smoke and Sulphur Dioxide* (Department of Trade and Industry, Stevenage, 1962–80).

Vacuoles as storage compartments for nitrate in barley leaves

Enrico Martinoia, Urs Heck & Andres Wiemken

Department of General Botany, Swiss Federal Institute of Technology, Sonneggstrasse 5, CH-8092 Zürich, Switzerland

Nitrate, the principal nitrogen source of most plants, can accumulate in large quantities in certain crop plants, notably members of the Chenopodiaceae (spinach and beet), Gramineae, Cruciferae (radish and kale) and Compositae (lettuce). Concentrations may exceed 2% fresh weight (17–24% dry weight) in extreme physiological conditions[1]. This is alarming because nitrate is readily reduced in organisms to the toxic nitrite, which may react with amines to form very potent carcinogenic nitrosamines[2]. Large stores of nitrate can be maintained in plant cells even in the presence of high nitrate reductase activity[3–8]. Incoming nitrate does not seem to mix with existing stores[4–6,9] and a steady influx of nitrate into the cells was found to be necessary to keep nitrate reductase stably induced[3,6]. Such observations have been explained by postulating a small 'metabolic' pool of nitrate accessible to nitrate reductase[4] and presumably also responsible for the induction of the enzyme[3], and a large 'storage' pool separate from the sites of metabolism[3–9,14]. Repeated attempts to measure these pools by an indirect method[4] have given equivocal results[1,10], but storage pools of nitrate have been generally thought to be located in vacuoles, which are difficult to isolate and analyse. However, we were able to isolate and purify the large central vacuoles of barley mesophyll cells and found most of the accumulated nitrate in the vacuoles.

Vacuoles were isolated from protoplasts as described in Fig. 1 legend and purity was confirmed by phase contrast microscopy (Fig. 1*c*). Contamination with intact protoplasts never exceeded 5%. Purity was further tested using marker enzymes. The specific activity of α-mannosidase[11,12] was much higher in the vacuolar fractions than in the homogenate of whole protoplasts (Table 1). Activities of α-mannosidase measured per 10^6 protoplasts and per 10^6 vacuoles were about the same (Table 1), indicating that almost all the enzyme was located in the vacuoles, of which each protoplast has only one (Fig. 1*b*). The vacuolar fraction contained 13% of the total activity of α-mannosidase present in the homogenate of protoplasts. On a basis of number protoplasts lysed the yield of α-mannosidase (that is of vacuoles) was 26%. The activities of two other hydrolytic enzymes, β-*N*-acetylglucosaminidase and protease, were also mostly restricted to the vacuoles (Table 2), supporting results obtained with other plants[11–13,15]. However, enzymes attributed to the chloroplasts (aldolase) and the cytosol (glucose-6-P isomerase, nitrate reductase[16]) were hardly detectable in the vacuolar preparation (Table 2).

We measured the contents of nitrate and amino acids in our pure preparation of vacuoles. As soluble substances can be easily lost from the vacuoles by leakage during isolation, we were surprised to find that almost all the nitrate in the protoplasts was contained in the vacuoles, together with about half of the soluble amino acids (Table 2). Vacuolar pools of amino acids have been reported elsewhere[17–19].

This demonstration of the vacuolar localization of nitrate raises the question of the value of nitrate storage pools to the plants. These pools could be nitrogen reserves. The rapid accumulation of nitrate from the soil could also help to prevent

Table 1 α-Mannosidase as vacuolar marker

	α-Mannosidase activity (nmol min^{-1})	
	Protoplasts	Vacuoles
Per (mg) protein	6.8	138*
Per 10^6 protoplasts or vacuoles	8.6	8.3†

* The specific activity of α-mannosidase in the isolated vacuoles is 20.3 times higher than that in the homogenate of protoplasts.

† Assuming that each protoplast has only one vacuole, 97% (±12%) α-mannosidase was found in the vacuoles. Procedures described in Table 2 legend.

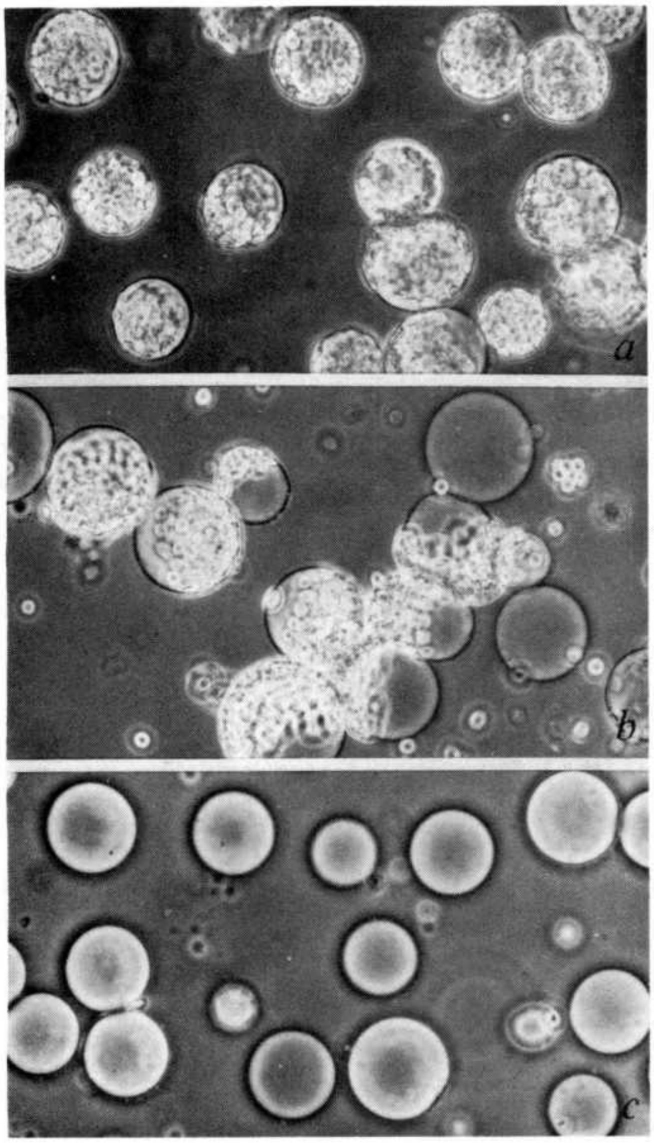


Fig. 1 *a*, Barley mesophyll protoplasts; *b*, protoplasts releasing their single large vacuole; *c*, isolated vacuoles. Phase contrast micrographs × 300. Protoplasts were prepared from the mesophyll of primary leaves of barley seedlings (*Hordeum vulgare* L. cv Nymphe) grown for 8 days on vermiculite, irrigated with 30mM KNO_3 (27 °C, 12 h light/dark cycle, 60% humidity). The leaves were cut into pieces (1 mm^2) which were incubated in medium containing 2% Onozuka R-10 cellulase (supplied by Welding & Co. Hamburg, Germany), 1‰ Pectolyase (Seishin Pharma, Noda, Chiba, Japan), 0.4M sucrose, 20mM citrate-Na^+ buffer, *p*H 5.6 (2 g fresh weight of leaves per 10 ml in a Petri dish). After 3 h at 20 °C the protoplasts, which came only from the mesophyll, were filtered through a 80 μm nylon net and purified by flotation (7 min at 1,400*g* in a swingout rotor) through a three-step density gradient: (step 1), 10 ml suspension of protoplasts diluted with 10 ml of 10% Ficoll 400 (Pharmacia Uppsala, Sweden), 0.4M sucrose, and 'supplements' (15mM phosphate buffer (*p*H 7.6), 2mM EDTA-Na (but not for nitrate determination), 10mM cystein-HCl (only in nitrate reductase determination)); (step 2), 10 ml 2.5% Ficoll, 0.4M sucrose and 'supplements'; (step 3), 4 ml 0.4M sorbitol and 'supplements'. The protoplasts recovered after centrifugation from the interface above step 2 were diluted with an equal volume of 10% Ficoll, 0.4M sucrose plus 'supplements' and flotated a second time through an identical density gradient. Vacuoles were liberated from the purified protoplasts by pressing the suspension (3×10^6 protoplasts per ml, 4 °C) through a long needle (10 cm × 0.2 mm) using a syringe. After this treatment about half of the protoplasts were ruptured and about two thirds of these released their single large vacuole intact. The vacuoles were immediately purified by flotation (2 min at 200*g*, followed by 3 min at 1,000*g* in a swingout rotor) through a three-step density gradient: (step 1), 5 ml lysate in 0.4M sucrose, 2.5% Ficoll and 'supplements'; (step 2), 5 ml 0.2M sucrose, 0.2M sorbitol and 'supplements'; (step 3), 2 ml 0.4M sorbitol and 'supplements'. Purified vacuoles were recovered from the step 3 layer. The whole isolation procedure from the time of rupturing of the protoplasts, was no longer than 8 min.

Table 2 Subcellular localization of enzymes and substances

	Activities or contents		
	Per 10^6 protoplasts	Per 10^6 vacuoles	% In vacuoles
Vacuolar marker enzymes:			
α-Mannosidase*	8.6	8.3	100
β-*N*-Acetylglucosaminidase*	12.8	11.0	89
Protease†	5.1	4.2	85
Extravacuolar marker enzymes:			
Aldolase*	82	0.3	<1
Malate dehydrogenase*	489	15.0	3
Glucose-P isomerase*	130	6.4	5
Nitrate reductase‡	40	<0.5	<1
Protein§	1.27	0.06	4.7
Amino acids‖	1,165	585	52
Nitrate‖	1,460	1402	99

Protoplasts were counted in a haemocytometer. The number of vacuoles was calculated from the α-mannosidase activity (the activity of α-mannosidase per 10^6 vacuoles (Table 1) was known and remained fairly constant (±8%) from one experiment to another). Per cent in vacuoles was calculated by assuming that 100% of the α-mannosidase is located in the vacuoles. The calculation was as follows: % activity or content in the vacuolar faction divided by % activity of α-mannosidase in the same fraction multiplied by 100. Enzyme activities were measured by standard methods: aldolase[22], malate dehydrogenase[22], nitrate reductase[23], α-mannosidase and *N*-acetylglucosaminidase with the corresponding nitrophenol substrates[12], and acid protease with haemoglobin as substrate in 0.1M sodium acetate buffer, *p*H 3.9, as described elsewhere[11]. Amino acids were measured with ninhydrin and protein by a modified method of Lowry[11]. Nitrate concentrations were determined potentiometrically with an ion-selective liquid membrane electrode as described elsewhere[24,25]. The membrane composition was 6% (by weight) methyl-tridodecyl-ammonium-chloride, 65% dibutylphthalate and 29% polyvinylchloride (by weight). The calibration for measurements in solution with protoplast lysate were done in solutions of the composition given in Fig. 1 legend (step 2) and in 10^{-4}, 10^{-3} and 10^{-2} M sodium nitrate. For the determination of nitrate in vacuoles the solution of step 3 was used as background. The detection limit for this experiment was checked using barley which had not been irrigated with nitrate.

* nmol min^{-1}.

† μg bovine serum albumin (BSA) equivalents per min.

‡ nmol NO_2^- h^{-1}.

§ μg BSA equivalents.

‖ nmol.

leakage, or uptake by neighbouring plants. It may also play a part in ionic and osmotic balance of the cells. This is particularly likely in plants in which nitrate reaches high concentrations of up to 300 mM (ref. 1) (generally halophytes) and in which nitrate may be accumulated as an alternative to chloride[20]. The answer to the question on the control of transport of nitrate through the vacuolar membrane could shed light on the relationship between the nitrate pools and the highly complex regulation of nitrate reductase activity[21].

We thank D. Ammann and P. Anker for help with nitrate measurements and S. Türler for reading, and D. Furrer for typing the manuscript.

Received 1 September; accepted 6 November 1980.

1. Hewitt, E. J. Hucklesby, D. P., Mann, A. F., Notton, B. A. & Rucklidge, G. J. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 255–287 (Academic, London, 1979).
2. Walters, C. L. & Walker, R. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 637–648 (Academic, London, 1979).
3. Heimer, J. & Filner, P. *Biochim. biophys. Acta* **230,** 362–372 (1971).
4. Ferrari, E. T., Yoder, O. C. & Filner, P. *Pl. Physiol.* **51,** 423–431 (1973).
5. Aslam, M., Oaks, A. & Huffaker, R. C. *Pl. Physiol.* **58,** 588–591 (1976).
6. Shaner, D. L. & Boyer, J. S. *Pl. Physiol.* **58,** 499–504 (1976).
7. Jones, R. W. & Sheard, R. W. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 521–539 (Academic, London, 1979).
8. Ashley, D. A., Jackson, W. A. & Volk, R. J. *Pl. Physiol.* **55,** 1102–1106 (1975).
9. Martin, P. *Z. Pflanzenphysiol.* **70,** 158–165 (1973).
10. Hageman, R. H., Reed, A. J. Femmer, R. A., Sherrard, J. H. & Dalling, M. J. *Pl. Physiol.* **65,** 27–32 (1980).
11. Wiemken, A., Schellenberg, M. & Urech, L. *Archs Microbiol.* **123,** 23–35 (1979).
12. Boller, Th. & Kende, H. *Pl. Physiol.* **63,** 1123–1132 (1979).
13. Nishimura, M. & Beevers, H. *Pl. Physiol.* **62,** 44–48 (1978).
14. Oaks, A. & Bidwell, R. G. S. *An. Rev. Pl. Physiol.* **21,** 43–66 (1970).
15. Matile, P. *An. Rev. Pl. Physiol.* **29,** 193–213 (1978).
16. Wallsgrove, R. M., Lea, P. J. & Milfin, B. J. *Pl. Physiol.* **63,** 232–236 (1979).
17. Wiemken, A. & Dürr, M. *Archs Microbiol.* **101,** 45–57 (1974).
18. Dürr, M., Urech, K., Boller, Th., Wiemken, A., Schwencke, J. & Nagy, M. *Archs Microbiol.* **121,** 169–175 (1979).

19. Sasse, F., Backs-Hüsemann, D. & Barz, W. *Z. Naturf.* **34c**, 848–853 (1979).
20. Stewart, G. R. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 651–652 (Academic, London, 1979).
21. Srivastava, H. S. *Phytochem.* **19**, 725–733 (1980).
22. Bergmeyer, H. U. *Methods of Enzymatic Analysis* 2nd edn (Verlag and Academic, New York and London, 1974).
23. Hageman, R. H. & Hucklesby, D. P. *Meth. Enzym.* **23**, 491–503 (1971).
24. Hartman, K. *et al. Mikrochim. Acta* **II**, 235–246 (1978).
25. Nielsen, H. J. & Hansen, E. H. *Analytica chim. Acta* **85**, 1–16 (1976).

Subpopulations of cultured chick sympathetic neurones differ in their requirements for survival factors

David Edgar, Yves-Alain Barde & Hans Thoenen

Abteilung Neurochemie, Max-Planck-Institut für Psychiatrie,
D-8033 Martinsried, FRG

Peripheral autonomic and sensory neurones derived from the neural crest will survive *in vitro* only if the culture medium is supplemented with specific factors (for review see ref. 1). For example, the nerve growth factor (NGF), although supporting the survival of sympathetic and spinal sensory neurones[2], is ineffective on parasympathetic neurones[3], whereas medium conditioned by chick heart cells (HCM) supports the survival of all three neuronal types[4]. We showed previously that the requirements of cultured spinal sensory ganglion neurones for survival factors changed during development, and so provided a basis for the classification of such factors[5]. We now demonstrate that cultured post-mitotic neurones from chick paravertebral sympathetic ganglia respond differentially to NGF, HCM and medium conditioned by C6 glioma cells (GCM). Thus, these three survival factors are functionally distinct in that they support the survival in culture of discrete subpopulations of sympathetic neurones. Those subpopulations responding to HCM and NGF are shown to differ not only in their requirements for survival factors but also in their contents of the cholinergic and adrenergic marker enzymes choline acetyltransferase (ChAT) and tyrosine hydroxylase (TH).

Cultures were established using methods similar to those previously described[5,6], the age of the chick embryos being checked by staging according to the criteria of Hamburger and Hamilton[7]. Lumbar and sacral paravertebral sympathetic ganglia were taken from embryos incubated between 8 and 18 days, dissociated and the cells pre-plated to remove most of the non-neuronal cells[5,6]. The few remaining non-neuronal cells in the cultures were easily distinguishable from neurones, the latter having compact, round cell bodies with thin neurite processes at least five cell diameters in length. NGF, GCM and anti-NGF antibodies were prepared as previously described, GCM being used in the present experiments after 10-fold concentration[5]. HCM was prepared by the method of Collins[8], except that 5% heat-inactivated horse serum was substituted for fetal calf serum to cultivate the heart cells, and the HCM was concentrated two-fold by ultrafiltration (Amicon UM 10) before use.

Table 1 Survival of cultured 12-day chick sympathetic neurones

	% Neuronal survival in response to:			
	Factor shown	Double concentration of factor	Factor plus NGF (5 ng ml^{-1})	Factor plus HCM
Control medium	0.09 ± 0.10			
NGF (5 ng ml^{-1})	37.1 ± 4.5	39.0 ± 5.0		
HCM	45.0 ± 3.0	41.0 ± 6.0	88 ± 9.5	
GCM	6.1 ± 2.1	8.0 ± 2.5	55 ± 4.5	71 ± 7.8

Mean values ± s.e.m. are given ($n = 3$). Sympathetic neurones were taken from paravertebral ganglia of 12-day old embryonic chicks. After trypsinization and dissociation, the cells were pre-plated in 10-cm plastic tissue culture dishes for 150 min as previously described[5,6]. The culture medium was F14 (ref. 14), supplemented with 10% heat-inactivated horse serum, penicillin (100 U ml^{-1}) and streptomycin (100 μg ml^{-1}). The neuronal suspension was then removed, those cells excluding trypan blue (>90%) counted using a haemocytometer and plated in poly-DL-ornithine-coated dishes[8] at a density of 5,000–8,000 cells per 35-mm dish (equivalent to one ganglion) in 1.5 ml medium containing the factors shown. Where NGF was not added, the media were supplemented with 500 ng ml^{-1} anti-NGF antibodies to inhibit any endogenous NGF, as found, for example, in GCM[5]. Neurones were counted after 48 h culture, the number being checked to remain constant on two subsequent days.

Table 2 Enzyme levels of 12-day chick sympathetic neurones

	Specific activity (pmol per min per mg) ± s.e.m. ($n = 3$)		
	ChAT	TH	Ratio ChAT/TH
After 48 h culture with:			
NGF	102 ± 44	281 ± 75	0.36
HCM	842 ± 77	137 ± 31	6.07
NGF + HCM	855 ± 45	310 ± 40	2.76
After 48 h culture with NGF, followed by 48 h with both NGF and HCM	85 ± 30	210 ± 40	0.40
Paravertebral sympathetic chains	680 ± 28	266 ± 20	2.56

Cultures were set up as described in Table 1 legend, the number of cultured neurones being increased to 50,000–80,000. After culture for 2 days, neurones were collected and ChAT activity was measured[15] using the ChAT inhibitor naphthylvinyl pyridine[16] (50 μM) to confirm specificity of the assay. Tyrosine hydroxylase was measured by a modification of the method of Levitt *et al.*[17], the difference being that a tetrahydrobiopterin cofactor (640 μM) was used[18], in the presence of 20 mM ascorbate, 10^{-4} M ferrous ammonium sulphate and 24,000 U ml^{-1} catalase[19]. The tyrosine concentration was 23 μM and specificity of the assay was checked using the TH inhibitor α-methyl-tyrosine (10^{-4} M)[20]. Initial reaction rates were linear with time and amount of sample. Protein was estimated by the method of Bradford[21] using bovine γ-globulins as the standard.

After removal of most of the non-neuronal cells by pre-plating and in the absence of added factors, less than 0.1% of the sympathetic neurones taken from 12-day old embryos survived after 2 days in culture (Table 1). Addition of NGF, HCM or GCM at the start of the culture period resulted in the survival of differing proportions of neurones, although each of the three factors was shown to be at saturating concentration: doubling the concentration of a given factor did not increase the number of surviving neurones. Culture with a combination of factors, however, resulted in an additive effect (Table 1), a maximum of 90% 12-day old neurones surviving in the presence of both NGF and HCM. Because such a high survival is possible in culture, and because there was no great loss of cells during the dissociation procedure which led to a complete disruption of the ganglia, this system provides an opportunity to observe *in vitro* a representative picture of the whole neurone population present *in vivo*. The differing proportions of neurones surviving in response to the different factors and the additivity of the effects of those factors point to the presence of discrete subpopulations of sympathetic neurones, the survival in culture of each being selectively supported by a given factor. This conclusion supports and extends the previous observation that not all chick sympathetic neurones respond to NGF in cultured ganglion explants[9].

To analyse further the relationships between the survival factors and their responsive neurones, we examined the effects of NGF, HCM and GCM on the survival of neurones taken from sympathetic ganglia of different embryonic ages. Figure 1 shows

0028-0836/81/040294—02$01.00

that NGF supports the survival of a maximum of 37% cultured neurones taken from 12-day old embryos, whereas HCM is maximally active on 14-day old neurones (57% survival) and the effect of GCM increases with age throughout the time period studied. The developmental patterns of sympathetic neuronal survival in response to NGF and GCM are strikingly similar to those previously obtained for dorsal root ganglion neurones[5]. Thus, NGF and GCM act as functionally distinct factors for both sensory and sympathetic neurones, and HCM may be differentiated from the other two on the basis of the age of the neurones on which it is maximally active. We have not demonstrated whether those neurones which are maximally dependent on NGF for their survival at day 12 become dependent on either HCM or GCM later during development. Furthermore, a relationship remains to be established between the various subpopulations defined here by their factor requirements in culture, and those subpopulations defined by morphological criteria *in situ*[10].

To approach these problems, we characterized biochemically the two major subpopulations of 12-day old neurones which survived in culture in response to either HCM or NGF. Table 2 shows that the HCM- and NGF-responsive neurones differ in their contents of TH and ChAT, necessary for the synthesis of the sympathetic neurotransmitters noradrenaline and acetylcholine, respectively. Thus, the cultured HCM-responsive neurones have a ChAT/TH ratio of 6.1, some 17-fold greater than that of those neurones surviving with NGF. Cells cultured with both NGF and HCM have an intermediate ChAT/TH ratio of 2.8, similar to that determined for freshly dissected 12-day old sympathetic chains.

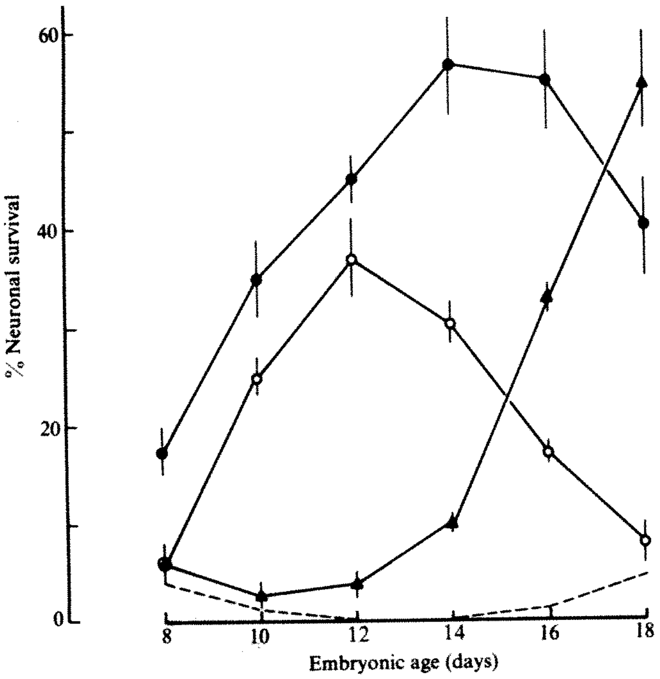


Fig. 1 Neurones were dissociated from paravertebral sympathetic ganglia taken from chicks of embryonic ages shown. After pre-plating, 5,000–8,000 neurones were plated on poly-DL-ornithine as described in Table 1 legend. Neuronal survival was estimated between days 2 and 5 after the start of culture, and it was checked that the percentage neuronal survival did not decrease significantly on two subsequent days. The dashed line shows the low basal levels of neuronal survival without added factors. An increased neuronal survival seen in control cultures of neurones from 8- and 18-day old chicks is a consequence of the slower death rate of the younger cells, and of the presence of more non-neuronal cells which support neurones in the cultures from older embryos. Parallel cultures were set up with supplements of: 5 ng ml^{-1} NGF (○), HCM plus 500 ng ml^{-1} NGF antibodies (●) and GCM plus 500 ng ml^{-1} NGF antibodies (▲). Results are expressed as percentages of surviving neurones ± s.e.m. ($n = 3$).

We do not know whether the high ChAT levels found in the HCM-responsive neurones is a consequence of a selection of those cells in culture which had pre-existing high ChAT activity *in vivo* or whether this enzyme was induced in the chick neurones by HCM in a manner similar to that shown for neonatal rat sympathetic neurones (for review see ref. 11). Note in this connection that HCM maintains the ChAT activity in cultured neurones at the levels measured in the embryonic sympathetic chains (Table 2, see also ref. 12) at a time when preganglionic (cholinergic) innervation is largely undeveloped[13], that is, when most of the ChAT may be in ganglionic neurones rather than innervating axons. Furthermore, HCM does not induce ChAT activity in that subpopulation of neurones previously selected by culture with NGF (Table 2). These observations are consistent with the hypothesis that NGF and HCM selectively support the survival of distinct subpopulations of embryonic chick sympathetic neurones which differ in their contents of adrenergic and cholinergic marker enzymes.

We thank Veronika Böhm for technical assistance.

Received 26 August; accepted 19 November 1980.

1. Varon, S. S. & Bunge, R. P. *A. Rev. Neurosci.* **1,** 327–361 (1978).
2. Levi-Montalcini, R. & Angeletti, P. U. *Devl Biol.* **7,** 653–659 (1963).
3. Helfand, S. L., Smith, G. A. & Wessels, N. K. *Devl Biol.* **50,** 541–547 (1976).
4. Helfand, S. L., Riopelle, R. T. & Wessels, N. K. *Expl Cell Res.* **113,** 39–45 (1978).
5. Barde, Y. A., Edgar, D. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1199–1203 (1980).
6. Greene, L. A. *Devl Biol.* **58,** 96–105 (1977).
7. Hamburger, V. & Hamilton, H. L. *J. Morph.* **88,** 49–92 (1951).
8. Collins, F. *Devl Biol.* **65,** 50–57 (1978).
9. Chamley, J. H., Mark, G. E., Campbell, G. R. & Burnstock, G. *Z. Zellforsch.* **135,** 287–314 (1972).
10. Edds, L. L. & van Horn, C. *J. comp. Neurol.* **191,** 65–76 (1980).
11. Patterson, P. H. *A. Rev. Neurosci.* **1,** 1–17 (1978).
12. Marchisio, P. C. & Consolo, S. *J. Neurochem.* **15,** 759–764 (1968).
13. Ross, L. L., Cosio, L., Witczak, L. & Smolen, A. *Anat. Rec.* **190,** 525 (1978).
14. Vogel, Z., Sytkowski, A. J. & Nirenberg, M. W. *Proc. natn. Acad. Sci. U.S.A.* **69,** 3180–3184 (1972).
15. Fonnum, F. *J. Neurochem.* **24,** 407–409 (1975).
16. White, H. L. & Cavallito, C. J. *J. Neurochem.* **17,** 1579–1589 (1970).
17. Levitt, M., Gibb, J. W., Daly, J. W., Lipton, M. & Udenfriend, S. *Biochem. Pharmac.* **16,** 1313–1321 (1967).
18. Black, I. B. *Brain Res.* **95,** 170–176 (1975).
19. Lerner, P., Nosé, P., Ames, M. P. & Lovenberg, W. *Neurochem. Res.* **3,** 641–651 (1978).
20. Udenfriend, S., Nirenberg, P. Z. & Nagatsu, T. *Biochem. Pharmac.* **14,** 837–845 (1965).
21. Bradford, M. M. *Analyt. Biochem.* **72,** 248–254 (1976).

The number of polarizing region cells required to specify additional digits in the developing chick wing

C. Tickle

Department of Biology as Applied to Medicine, The Middlesex Hospital Medical School, Cleveland Street, London W1P 6DB, UK

A group of mesenchyme cells at the posterior margin of the developing wing—the polarizing region—can have a dramatic effect on the pattern of structures which develop across the antero-posterior axis of the limb. If one grafts a polarizing region to the anterior margin of a wing bud so that this bud now has one polarizing region at the anterior margin and one at the posterior margin, the wing that develops has duplicated structures across the antero-posterior axis in mirror-image symmetry[1-3]. That is, the normal pattern of digits 2 3 4 (reading from anterior to posterior) becomes 4 3 2 2 3 4. The ability of the polarizing region to specify additional digits from adjacent tissue can be progressively attenuated by γ-ray radiation[4]. If this attenuation is caused by progressively fewer cells remaining viable to signal, there should be a quantitative relationship between the number of polarizing cells used and the digit specified next to the graft. Here, two tests are reported which confirm this idea. The results suggest that the apical ridge cooperates with a small number of polarizing region cells in a monolayer to specify structures across the antero-posterior axis of the wing.

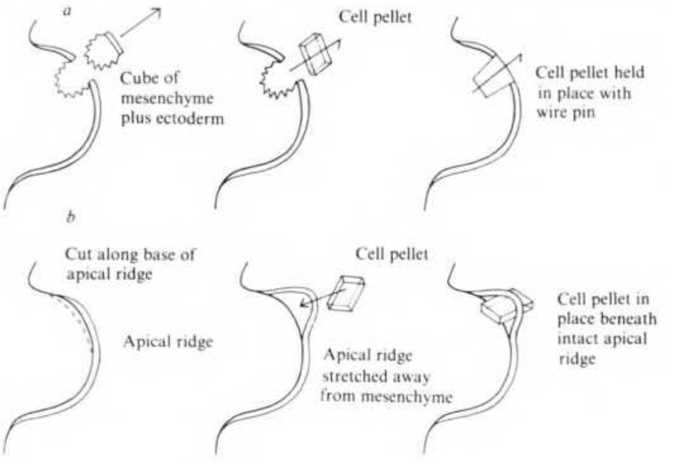

a
Cube of mesenchyme plus ectoderm
Cell pellet
Cell pellet held in place with wire pin
b
Cut along base of apical ridge
Apical ridge
Cell pellet
Apical ridge stretched away from mesenchyme
Cell pellet in place beneath intact apical ridge

Fig. 1 Diagram to show how grafts of pieces of cell pellet were made to host wing buds (stage 20–21). The grafts were cut from a centrifuged pellet of cells which had been incubated for 1 h at 37 °C. The grafts were ~150–200 μm per side and contained about 6,000 cells (estimated from serial sections: average for 4 pieces = 5,979, s.d. 464). The cell suspensions were prepared from polarizing regions and equivalent-sized pieces of anterior mesenchyme tissue dissected from the wing buds of chick embryos (stage 22: Hamilton–Hamburger) following removal of the ectoderm with 2% trypsin (Difco). Various mixtures of polarizing region cells and inactive cells from anterior margins were made. For each suspension a total of 20 pieces of tissue was used. The standard grafting procedure, shown in *a*, involved removing a piece of mesenchyme plus overlying apical ridge from the anterior margin of the host limb bud opposite somite 16 or 17 and placing the piece of pellet in the hole. The graft was held in place by a pin of platinum wire. In the other grafting procedure, shown in *b*, a cut was made along the base of the apical ectodermal ridge over the anterior part of the host bud with a sharpened tungsten needle. The apical ectodermal ridge was raised up from the underlying mesenchyme and the piece of pellet was placed between. No mesenchyme tissue was removed. The apical ridge was stretched away from the mesenchyme and once the graft was in place it contracted back to hold the pellet in position.

Table 1 Effect of diluting polarizing cells with anterior margin cells on the digit formed next to the graft

% Of polarizing cells in pellet	No. of cases	Extra digit adjacent to graft				
		4	3	2	Knob	None
a, Standard grafting procedure shown in Fig. 1*a*						
100	5	3	2	0	0	0
100 (50/50 wing/leg)	9	4	2	3	0	0
Total	14	7	4	3	0	0
55 (quail)	4	0	4	0	0	0
50	11	0	4	5	0	2
50 (quail)	6	0	1	3	2	0
Total	21	0	9	8	2	2
33	1	0	1	0	0	0
25	13	0	1	7	2	3
20	4	0	0	2	0	2
10	6	0	1	4	0	1
5	7	0	0	2	2	3
0	9	0	0	0	0	9
b, Grafts made to a slit beneath an intact apical ridge as shown in Fig. 1*b*						
100	11	11	0	0	0	0
50	7	2	4	1	0	0
20	3	0	3	0	0	0
9–10	13	0	7	6	0	0
0	4	0	0	0	0	4

In the first set of experiments, cells from the polarizing region were mixed with inactive cells from the region along the anterior margin of the wing bud. By progressively diluting disassociated polarizing region cells with these inactive cells and grafting the mixture after re-aggregation to an anterior position in another bud, a clear relationship between the number of polarizing region cells and the digit formed next to the graft can be shown.

The results of grafts carried out by the standard procedure (Fig. 1*a*), but with varying proportions of re-aggregated polarizing region cells and inactive anterior cells, are listed in Table 1*a*. As the proportion of polarizing region cells in the grafts decreased, the digit formed next to the graft changed progressively: first a digit 3 (Fig. 2*c*) instead of a digit 4 (Fig. 2*a*) was produced, then a digit 2 (Fig. 2*d*) instead of a digit 3, and finally no additional digits at all (Fig. 2*b*) were specified. A

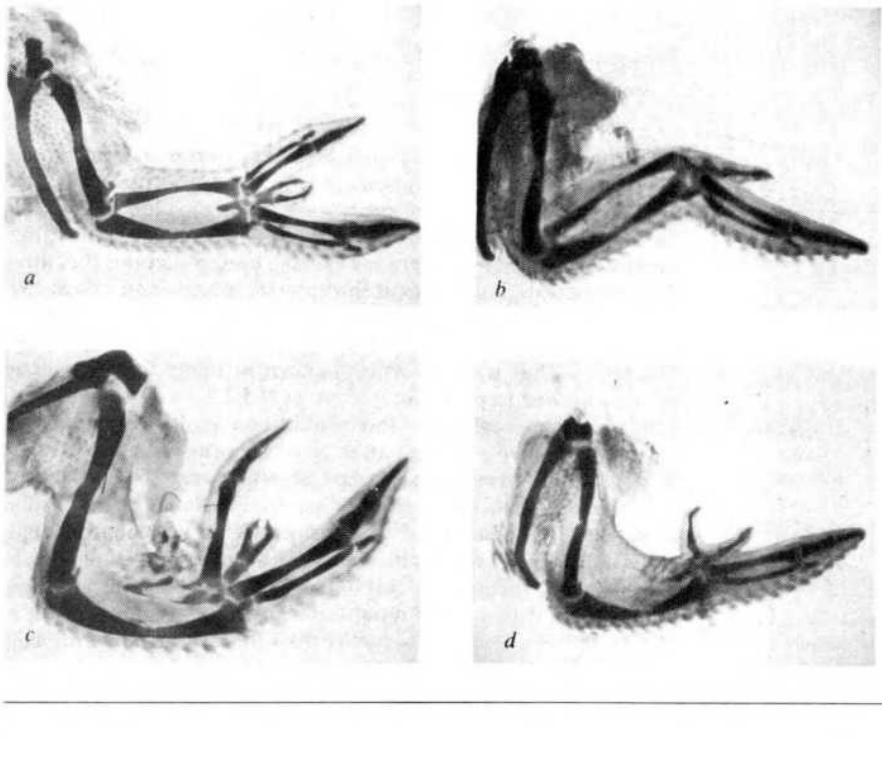


Fig. 2 Dorsal views of whole mounts of wings obtained following grafts of re-aggregated cells made as shown in Fig. 1*a*. *a*, 100% polarizing region cells. Pattern of digits is 4 3 2 2 3 4. *b*, 100% anterior margin cells. Pattern of digits is 2 3 4. *c*, 50% polarizing region cells and 50% anterior margin cells. Pattern of digits is 3 2 2 3 4. *d*, 5% polarizing region cells and 95% anterior margin cells. Pattern of digits is 2 2 3 4.

Table 2 Comparison of independent estimates of the number of polarizing region cells, placed under an intact apical ridge, which are required to specify each additional wing digit

	Digit 4	Digit 3	Digit 2
a	100(1)	79(10)	35(12)
b	164	64–32	32 or less

a, Mean numbers of polarizing cells on pieces of plastic film grafted under the apical ridge with the cells facing the ridge; *b*, number of polarizing cells on a face of a piece of pellet of re-aggregated polarizing and anterior mesenchyme cells grafted under an intact apical ridge. In *b*, data were calculated from the average number of cells on a face of a pellet and % polarizing cells in mixed pellet required to specify each digit. Values in parentheses indicate number of cases.

similar effect has also been obtained when polarizing region cells are diluted with BHK cells (unpublished observations). Even when only 5% of the cells were from the polarizing region, an additional digit 2 was obtained in one-third of the cases (Fig. 2*d*). Because the same order of loss of duplicate digits is found when the polarizing region is treated with increasing doses of γ-ray radiation[4], it seems likely that the effect of the irradiation is to diminish the number of viable cells in the graft.

To investigate whether close contact with the apical ridge leads to more efficient signalling, as suggested by the recent work of Saunders[5], another series of grafts, in which the polarizing region cells were progressively diluted with anterior mesenchyme cells, was carried out (Table 1*b*). In this series, however, the apical ridge was left intact (Fig. 1*b*). It is clear that these grafts placed under an intact apical ridge are much more effective. For example, if the effects of grafts containing 50% polarizing region cells are compared, digit 4 was still specified in some cases when grafts were placed under the ridge, whereas an extra digit 4 has never been obtained with a standard graft.

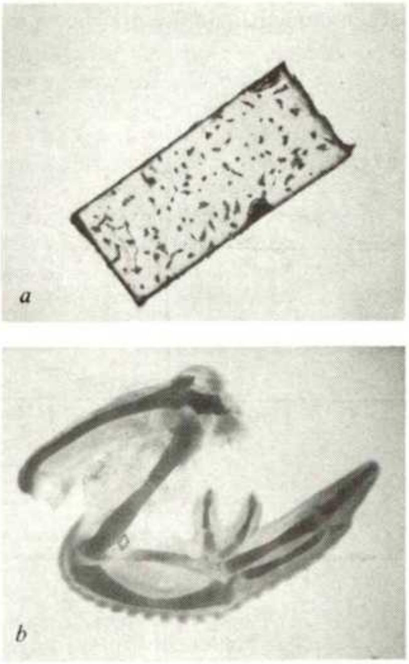


Fig. 3 *a*, Piece of plastic film (Melinex, ICI) after 3 h incubation on Parafilm (American Can Company) with a drop of suspension of polarizing cells prepared from at least eight polarizing regions. This piece of plastic (scale bar is 100 μm) has been fixed and stained with Leishman's to show attached cells, some of which have spread. To count the numbers of cells on such pieces of plastic (100–250 μm on a side) before grafting, the pieces were transferred with a spatula to a dish and viewed in an inverted phase-contrast microscope. The piece of plastic with a known number of polarizing cells was then carried into position over the host wing bud on a spatula and manoeuvred into the grafting site. *b*, Wing which developed after grafting of a piece of plastic film carrying 28 polarizing cells. The graft was made to a slit beneath the apical ridge and the film was oriented so that the cells faced the ridge. Digit pattern is 2 2 3 4. Note that the piece of plastic has ended up at the elbow of the developed wing.

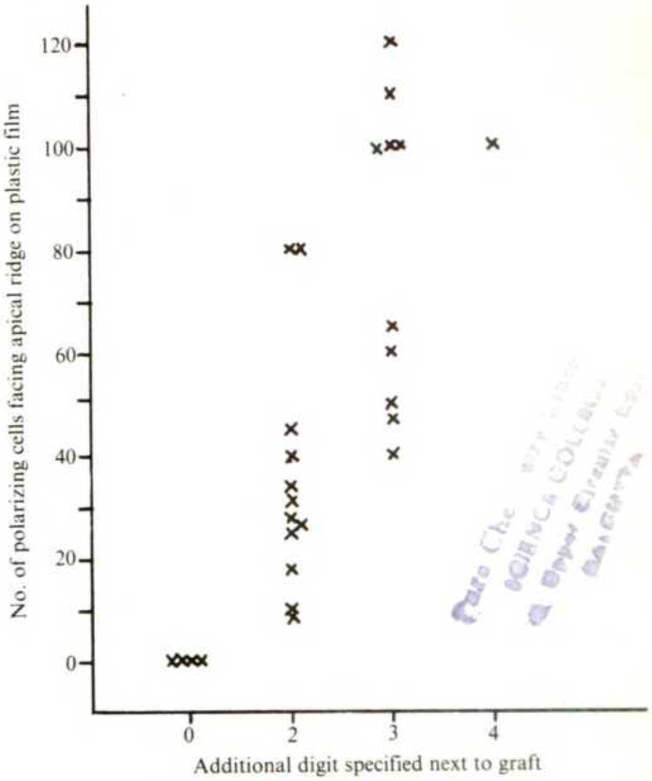


Fig. 4 Data on the digit which formed next to the graft after implantation of pieces of plastic film carrying varying numbers of polarizing cells. The plastic film was grafted to a slit beneath the apical ridge with the cells facing the ridge.

From the total numbers of cells in pieces of pellet the following estimates can be made: with the standard grafting procedure, at least 6,000 polarizing region cells (100% polarizing region cells in the graft) are required to obtain digit 4 next to the graft; to obtain digit 3, 3,000–6,000 cells are required, and for digit 2 between 3,000 and as few as 300 are needed. For grafts made under an intact apical ridge, a digit 4 requires about 3,000 polarizing region cells, but only about 600 cells are required for a digit 3 and even fewer for a digit 2.

In the second set of experiments, small numbers of polarizing cells stuck to pieces of plastic film (Fig. 3*a*) were implanted just beneath the apical ridge. Figure 4 shows that remarkably small numbers of polarizing cells can now specify additional digits in the chick wing (Fig. 3*b*). Grafts of pieces of plastic with small numbers of anterior mesenchyme cells (five cases) always resulted in normal wings. Figure 4 also shows a clear relationship between the number of polarizing cells grafted on the piece of film and the digit which developed next to the graft. The estimate of the mean number of cells required to specify each wing digit is shown in Table 2. Thus, to specify digit 4 at least 100 polarizing cells must be grafted beneath the apical ridge, whereas specification of digit 3 requires 79 cells and digit 2, 35 cells. In one case, the transfer of only nine polarizing cells led to a wing with an additional digit 2.

The two independent estimates of the number of polarizing cells in contact with the ridge that are required to specify each additional wing digit are apparently very different. However, if the number of polarizing cells on only the face of each mixed pellet, used for grafts under the apical ridge, are considered, the number of cells required for signalling agrees well with the estimates obtained from implanting known numbers of cells on pieces of plastic film (Table 2). Thus, it seems likely that the signalling ability of the polarizing region involves only a monolayer of cells, the rest of the cells in a grafted pellet being superfluous or inactive. Both this finding and the observation that the mixed aggregates signal best when grafted under an intact apical ridge suggest that cooperation between the monolayer of polarizing cells and the apical ectodermal ridge is essential for positional signalling.

This work was supported by the MRC. I thank Professors L. Wolpert and B. Alberts for their enthusiastic encouragement. I also thank Professor Alberts for two visits to his laboratory supported by NIH grant GM23928, and M. Goodman and A. Crawley for help with photography and histology.

Received 14 August, accepted 21 November 1980

1 Saunders, J W & Gasseling, M T in *Epithelial–Mesenchymal Interactions* (eds Fleischmajer, R & Billingham, R E) 78–97 (Williams & Wilkins, Baltimore, 1968)
2 Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199–202 (1975)
3 Tickle, C in *Development in Mammals* Vol 4 (ed Johnson, M H) 101–137 (North-Holland, Amsterdam, 1980).
4 Smith, J C, Tickle, C & Wolpert, L. *Nature* **272**, 612–613 (1978)
5 Saunders, J W in *Vertebrate Limb and Somite Morphogenesis* (eds Ede, D A, Hinchliffe, J. R & Balls, M) 1–24 (Cambridge University Press, 1977)

Flow cytometry analysis of T cells and continuous T-cell lines from autoimmune MRL/l mice

D. E. Lewis, J. V. Giorgi & Noel L. Warner

Immunobiology Laboratories, University of New Mexico School of Medicine, Albuquerque, New Mexico 87131

The study of spontaneous autoimmunity in mouse models affords an opportunity to determine the cellular basis of the immune dysregulation observed in this disease. Recently, a new mouse strain, MRL/Mp-*lpr*/*lpr* (MRL/l) has been developed[1] which carries an autosomal recessive gene (*lpr*) that results in massive lymph node enlargement concomitant with the development of several autoantibodies[1–4]. The interest in this strain lies in the possibility that the defect in T-cell regulation of the immune response is manifested at a different level from that in the NZB mouse[3]. It has been reported that the proliferating population of lymphoid cells in the nodes of these mice are T cells, but that many of them are devoid of Lyt surface antigens[4]. We have accordingly initiated several lines of research with these mice, including quantitative flow cytometry characterization of Lyt antigen expression of cells in the lymph nodes of the mice. In an approach to isolate and study the properties of these cells, we have also established continuous cell lines from the lymph node cells of MRL/l mice, using techniques similar to those used to establish continuous lines of antigen-activated cytotoxic T cells[5,6] and helper T-cell populations[7].

Lymph node cells from MRL/l and MRL/n mice at various ages, and T-cell lines which were established from MRL/l lymph node cells were examined for Lyt-1, Lyt-2 and Thy-1.2 cell surface density using indirect immunofluorescence. Xenogeneic (rat) hybridoma-derived monoclonal antibodies which have specificities for Lyt-1, Lyt-2 (both framework determinants) and Thy-1.2, were provided by Drs Ledbetter and Herzenberg of Stanford University[8]. The lymph node cells were stained for 20 min on ice, washed three times in Eisen's balanced salt solution (BSS) and then fluorescein-conjugated (Fl) goat anti-rat immunoglobulin was added as a second step reagent. The cells were incubated for an additional 20 min, washed three times in BSS and examined using the flow cytometry systems at Los Alamos Scientific Laboratories[9,10] or a Becton Dickinson FACS-III cell sorter.

Table 1 Comparison of cell-surface markers on *in vivo* MRL/lpr lymph node cells and *in vitro* MRL/lpr cell lines

Surface marker	MRL/lpr MLN cells	MRL/lpr T-cell line
Thy-1.2	Moderate density (lower than MRL/n)	Moderate density
Lyt-1	Increased % positive Lower density	Most positive Low density
Lyt-2	Marked reduction in % positive	Most lines variable
Lyt-3	As for Lyt-2	As for Lyt-2
T-30	Lower density	Most lines positive Low density
Th-B	Negative	Negative
FcR	Negative	Negative

Surface marker was detected by flow cytometry with monoclonal antibodies. Results for MRL/lpr are intensity and per cent cells relative to age matched control MRL/n mice. The MRL/lpr T-cell line is representative of most MRL/lpr cell lines. MLN, mesenteric lymph node.

Attempts were made to establish continuous cell lines from MRL/l, MRL/n and C57BL/6 mesenteric lymph node cells by plating 5×10^5 to 1×10^6 cells ml^{-1} in 5-ml volumes of Dulbecco's modified Eagle's medium (DMEM) supplemented with non-essential amino acids (1 ml 100 ml^{-1}, 100x, Gibco), 10^{-4} M 2-mercaptoethanol, 10% fetal calf serum and 25% T-cell growth factor (TCGF) (v/v), which was prepared from rat spleen cells stimulated with 5 μg ml^{-1} concanavalin A (Con A) for 46 h (ref. 5).

The results of the analysis of Lyt phenotypes of lymph node cells from MRL/l and MRL/n mice are shown in Fig. 1. The

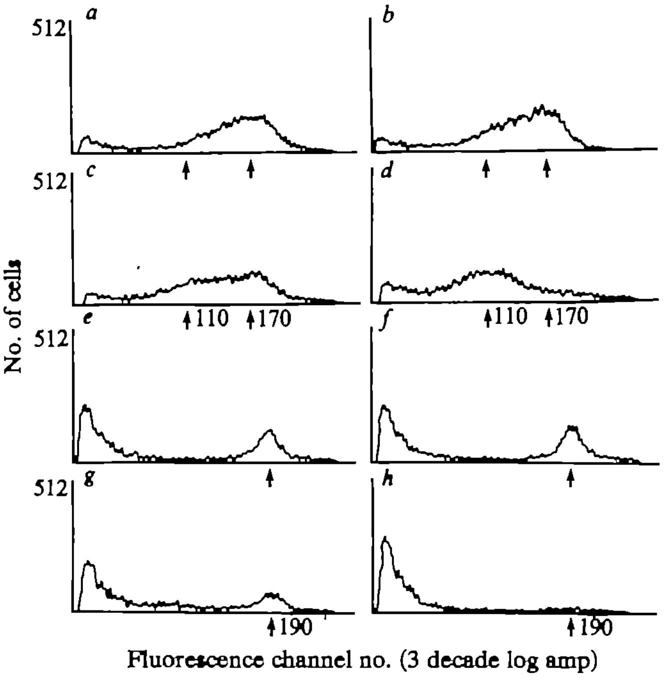
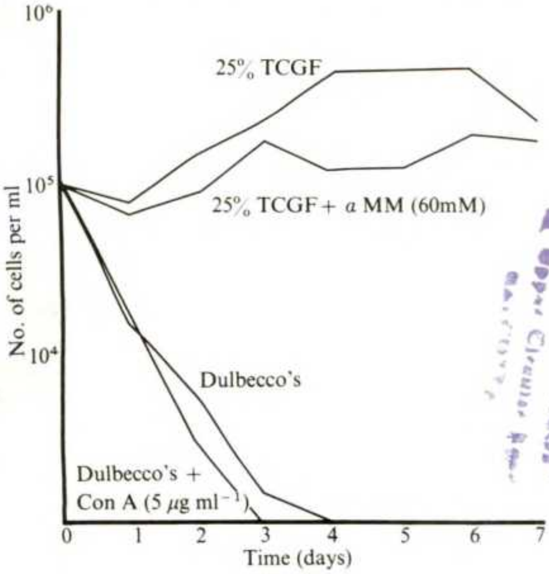


Fig. 1 Flow cytometry analysis of Lyt-1 and Lyt-2 expression on mesenteric lymph node cells of MRL/lpr and MRL/n mice. In each analysis cell suspensions were stained with the hybridoma-derived supernatants followed by fluorescein-conjugated SJL anti-rat immunoglobulin. The left column of panels (*a*, *c*, *e*, *g*,) shows cell suspensions from 2-month old mice, and the right-hand column of panels (*b*, *d*, *f*, *h*) those from 5-month old mice. *a*, *b*, MRL/n cells stained with anti-Lyt-1; *c*, *d*, MRL/lpr with anti-Lyt-1; *e*, *f*, MRL/n with anti-Lyt-2; *g*, *h*, MRL/lpr with anti-Lyt-2. The key distinction between MRL/lpr and MRL/n mice is the presence of a low density Lyt-1 peak around channel 110, and a selective reduction in Lyt-2 cells in older MRL/lpr mice

0028-0836/81/040298—03$01 00



Science College Upper Circular Road

Fig. 2 Dependence on T-cell growth factor (TCGF) of MRL-III continuous T-cell line. Four aliquots of MRL-III cells were seeded at $10^5\,\mathrm{ml}^{-1}$ in the presence of Dulbecco's medium with the indicated additive. Growth occurs in the presence of 25% TCGF either with or without α-methylmannoside. Growth of the line is not supported by the addition of Con A alone.

MRL/l lymph node population at 8 weeks of age is characterized by a shift in the surface density of Lyt-1, in that, compared with MRL/n lymph node cells, a population of cells with lower Lyt-1 expression is evident. By 5 months of age this low surface density population of Lyt-1 cells has become the predominant population and by 5 months of age there are very few Lyt-2-positive lymph node cells (that is, either Lyt-1^+2^+ or Lyt-1^-2^+). The Thy-1.2 staining characteristics indicate that the MRL/l mouse lymph node cells, which are all Thy-1.2 positive, have lower amounts of Thy-1.2 antigen than do MRL/n lymph node cells. As the cell population is also somewhat larger, the actual cell-surface density is considerably lower than in the MRL/n mice.

Seven MRL/l cell lines have now been established and several of the lines have been in culture for up to 8 months. The lines were all established from mesenteric lymph node cells from 3–5-month old MRL/l mice. The cells will not grow in DMEM without TCGF, nor will they grow in DMEM supplemented with $5\,\mu\mathrm{g\,ml}^{-1}$ Con A (Fig. 2). Furthermore, addition of 60 mM α-methylmannoside to the growth medium does not inhibit the effect of the TCGF. These data indicate that the growth of the MRL/l cell lines is dependent on a growth factor in the TCGF medium which is not Con A itself.

Attempts made to establish continuous cell lines from MRL/n and C57BL/6 lymph node cells were unsuccessful. The cells multiplied much more slowly than the cells from MRL/l mice, and never reached high cell densities *in vitro*. These lymph node cells were occasionally maintained in culture for periods of up to 2 months.

Because it is important to know whether these *in vitro* continuous T-cell lines are representative of the abnormal proliferating cells *in vivo*, we have compared the *in vivo* and *in vitro* cell lines using a range of surface markers. After several weeks in culture, the *in vitro* lines show a complete absence of B cells, and all cells are Thy-1.2 positive. The Lyt phenotype of the cells has fluctuated, however, and as shown in Fig. 3, includes both Lyt-1^+2^- and Lyt-1^+2^+ lines. The predominant phenotype of most of the lines has been Lyt-1^+2^-, with fluctuations primarily among cells that are Lyt-2^+. Some lines thus contain many Thy-1^+ Lyt-null cells. These may, however, have lower undetectable amounts of Lyt-1.

Using a broader range of cell-surface markers, we show in Table 1 that the *in vitro* lines, as studied here, are generally similar to the predominant population of cells *in vivo*.

Our results indicate that the massive lymphoproliferation observed in MRL/l mice is associated with a decline in the proportion of Lyt-2-positive T cells. This population includes both Lyt-123- and Lyt-23-positive cells. The majority of the cells present in these large nodes are Lyt-1 positive, but have a distinctly lower density of Lyt-1 on their surfaces than normal cells. These results are in contrast to those of Theofilopoulos *et al.*[4], who reported that about 50% of the cells in the MRL/l lymph node are Lyt-null T cells, although they did suggest that the 'Lyt-null cells' may not be completely devoid of Lyt antigens. With flow cytometry we find only approximately 10%

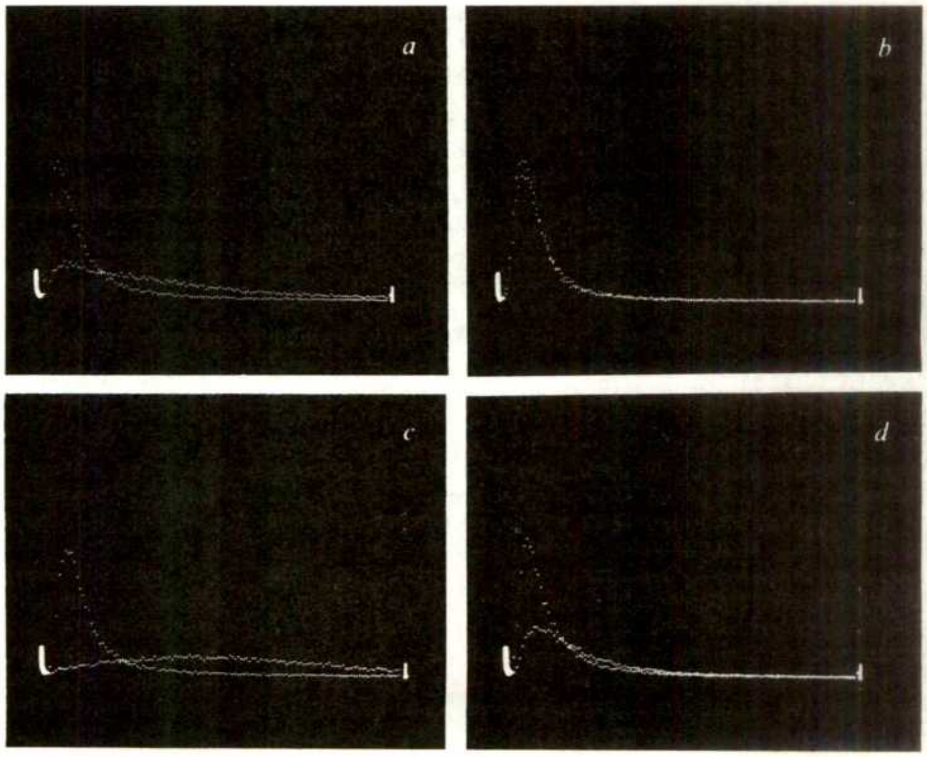


Fig. 3 Flow cytometry analysis of Lyt-1 and Lyt-2 expression on two MRL continuous T-cell lines. Each panel shows the fluorescence profile (FACS-III histogram) of the cell suspension stained with either medium or hybridoma supernatant, followed by fluorescein-conjugated goat anti-rat immunoglobulin. Panels *a* and *b* show cell line MRL-VI stained, respectively, with anti-Lyt-1 and anti-Lyt-2, and panels *c* and *d* cell line MRL-VII stained, respectively, with anti-Lyt-1 and anti-Lyt-2. In each panel the number of cells is on the ordinate and the relative fluorescent intensity on the abscissa.

of the T cells in the MRL/l lymph node to be of Lyt-null type (but Thy-1.2 positive). Thus, this quantitative analysis clearly indicates that the proliferating T-cell population in the MRL/l lymph node may be a relatively homogeneous T-cell subset. Further analysis of these cells with other differentiation antigens is proceeding.

The present report also extends the findings of other investigators which indicate that continuously proliferating lines of T cells can be readily established in long-term culture as long as T-cell growth factors are present. Our results indicate that it is possible to establish cells from MRL/l mice (but not cells from normal mice) in continuous culture without deliberate exogenous antigenic stimulation. Thus, it might be proposed that antigenic stimulation for these lymph node cells from MRL/l mice is from an autoantigenic source, probably derived *in vivo* before culture, as the lines do require continual TCGF for maintenance.

It is interesting that the cell lines generated from the lymph nodes of MRL/l mice are variable in their Lyt expression, although all are T cells. Cloning and cell sorting experiments are in progress to resolve this problem and to determine whether this variability truly represents instability in Lyt expression in a continuously proliferating clonal line. The fact that cell lines can be established at all in the absence of antigenic stimulation may allow the development of pure subpopulations of specific Lyt antigen-bearing cells for further dissection of the functional role of Lyt subsets in the aberrant immune response observed in these autoimmune mice.

This work was supported by NIH research grants AM19841 and fellowships AI05779 and CA05921. We thank Mrs Jerri Davis for maintenance of the MRL mouse colonies.

Received 7 May; accepted 24 October 1980.

1. Murphy, E. D. & Roths, J. B. *Fedn Proc.* **36,** 1246 (1977).
2. Murphy, E. D. & Roths, J. B. in *Genetic Control of Autoimmunity* (eds Rose, N. R., Bigazzi, P. E. & Warner, N. L.) 207–221 (Elsevier, New York, 1978).
3. Gershon, R. K., Horowitz, M., Kemp, J. D., Murphy, D. B. & Mruphy, E. D. in *Genetic Control of Autoimmunity* (eds Rose, N. R., Bigazzi, P. E. & Warner, N. L.) 223–227 (Elsevier, New York, 1978).
4. Theofilopoulos, A. N., Eisenberg, R. A., Bourdon, M., Crowell, J. S. Jr & Dixon, F. J. *J. exp. Med.* **149,** 516–534 (1979).
5. Gillis, S. & Smith, K. A. *Nature* **268,** 154–156 (1977).
6. Nabholz, M., Engers, H. D., Callovo, D. & North, M. *Curr. Topics Microbiol. Immun.* **81,** 176 (1978).
7. Watson, J. A., Aarden, L. A. & Lefkovitz, I. *J. Immun.* **122,** 209 (1979).
8. Dennert, G. *Nature* **277,** 476–477 (1979).
9. Shreier, M. H. & Tees, R. *Int. Archs Allergy appl. Immun.* **61,** 227–237 (1980).
10. Ledbetter, J. A. & Herzenberg, L. A. *Immun. Rev.* **47,** 63–90 (1979).
11. Steinkamp, J. A. *et al. Devs scient. Instrum.* **44,** 1301–1310 (1973).
12. Warner, N. L., Daley, M. J., Richey, J. & Spellman, C. *Immun. Rev.* **48,** 197–243 (1979).

Cyclosporin A promotes spontaneous outgrowth *in vitro* of Epstein–Barr virus-induced B-cell lines

A. Graham Bird*‡, Sandra M. McLachlan† & Sven Britton*

* Department of Immunobiology, Wallenberglaboratory,
Karolinska Institute, S-104 05 Stockholm, Sweden and
Department of Infectious Diseases, Danderyd Hospital,
S-182 03 Danderyd, Sweden
† Department of Medicine, University of Newcastle-upon-Tyne,
Newcastle-upon-Tyne NE1 7RU, UK

Cyclosporin A (CSA) is a fungal metabolite which exerts profound effects on the immune system and has potential as a selective immuno-suppressive agent[1]. Clinical trials with human renal allograft recipients have confirmed this potential but there have been disturbing reports of lymphoma in a significant number of patients[2,3]. Despite extensive animal studies, the specificity of this drug for human lymphocyte subpopulations is largely unknown[4]. We demonstrate here that *in vitro* CSA has no apparent effect on Epstein–Barr virus (EBV)-induced B-lymphocyte activation, but totally inhibits the T-cell dependent pokeweed mitogen (PWM) B-cell response. In addition, CSA markedly facilitates the outgrowth of B-lymphoblastoid cell lines from both EBV-infected and non-infected lymphocytes of EBV immune donors cultured *in vitro*. These results indicate that CSA can interfere with the lymphocytes normally responsible for maintaining the life-long carrier state initiated by primary infection with EBV, allowing outgrowth of the persistently infected cells circulating in the peripheral blood.

We have investigated the effects of CSA on human polyclonal B-cell activation *in vitro*. Figure 1 demonstrates the differential effects of CSA on activation induced by two mechanistically different activators, EBV and PWM. The results clearly show that CSA completely inhibits PWM-induced B-cell activation which is T-cell and monocyte dependent[5,6] whereas the polyclonal activation produced by the T- and monocyte-independent EBV[6] is actually enhanced early in the culture period at 6 days. CSA did not influence cell viability in the PWM cultures but markedly reduced morphological blast transformation.

We next studied the effects of CSA on the generation of lymphoblastoid cell lines from individuals of known EBV immune status whose lymphocytes were either infected *in vitro* with EBV or uninfected. Table 1 summarizes the results obtained from a number of individuals. In cultures of lymphocytes not treated with CSA, lymphoblastoid cell lines were only obtained from cultures to which EBV had been added. In agreement with the results of others[7,8], cell lines were easily obtained from lymphocytes from a seronegative donor infected

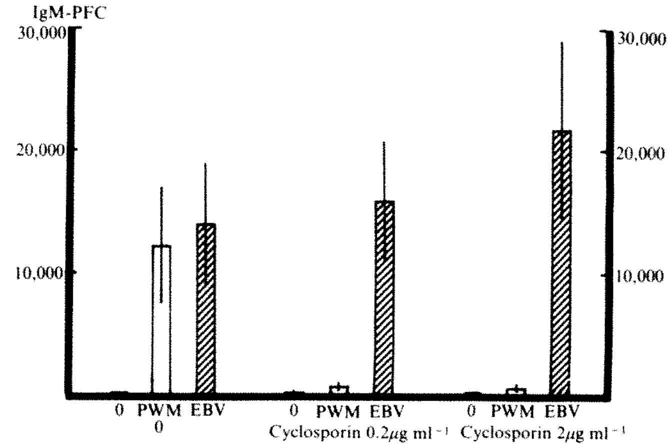


Fig. 1 The effects of cyclosporin A on the polyclonal activation of human blood lymphocytes by EBV and PWM. Lymphocytes from seven adult volunteers of known EBV immune status were cultured at a final density of 10^6 ml^{-1} in RPMI 1640 medium supplemented with 10% human AB serum (anti-VCA positive) and antibiotics. Cultures were incubated for 6 days in 50-ml plastic flasks (Falcon 3013) in the presence of 5% CO_2. EBV infection was achieved by 1 h pre-incubation at 37 °C of a suspended pellet of 5×10^6 lymphocytes with 1 ml of B95-8 cell line supernatant. PWM was added to appropriate cultures at a final concentration of 20 μg ml^{-1}. Cyclosporin A (Sandoz) was initially dissolved in a small volume of 95% ethanol and then diluted in Earle's balanced salt solution to give a concentration of 100 μg ml^{-1}. The drug was added to cultures to give a final concentration of 2.0 or 0.2 μg ml^{-1}. Cultures were assayed for the presence of antibody-secreting cells by use of the protein A plaque assay which has been described elsewhere[6]. Results are expressed as IgM plaque-forming cells (PFC) per 10^6 viable cells at the end of culture. A similar relation was found for IgG PFC (data not shown).

---

‡ To whom correspondence should be addressed at the Department of Immunology, The Medical School, Vincent Drive, Birmingham B15 2TJ, UK.

0028-0836/81/040300—02$01.00

Table 1 Effect of cyclosporin A on the *in vitro* outgrowth of lympoblastoid cell lines from individuals of known EBV-immune status

Donor	Serological status (anti-VCA)	CSA addition (2 μg mg^{-1})	EBV infection *in vitro*	Cell lines established
1–6	+	+	+	+
		–	+	–
		+	–	+
		–	–	–
7–8	+	+	+	+
		–	+	–
		+	–	–
		–	–	–
9	+	+	+	+
		–	+	+
		+	–	+
		–	–	–
10	–	+	+	+
		–	+	+
		+	–	–
		–	–	–

Immune status of the donors was examined by determination of the antibody titre to EB-viral capsid antigen (VCA). Cultures were established as in Fig. 1 and were regularly examined for cell line formation. This was determined visually by the appearance of progressively growing foci of transformed cells[7]. The scores above are from 4–12 weeks of culture. In many cases the presence of lines was confirmed by transferring the cells into secondary cultures. Cell lines arising in non-EBV-infected lymphocytes were examined for the presence of EBNA by the method of Reedman and Klein[9].

with EBV, but, despite initial transformation in all seropositive donor cells, regression of cell lines occurred in eight out of nine donor cultures. Addition of CSA allowed cell lines to be established by 4 weeks with ease and without regression from all seropositive donors so far tested. More significantly, B-cell lines have also been obtained from cultured lymphocytes from six out of nine seropositive donors exposed to CSA but not infected with EBV. These cell lines contained the EB-determined nuclear antigen (EBNA), a characteristic of EBV transformation[9] and they were of similar morphology to those derived from cells exposed to EBV (K. Nilson, personal communication).

These results clearly demonstrate that CSA exerts profoundly selective effects on human lymphocyte subpopulations *in vitro*. The enhancement of the polyclonal immunoglobulin response induced by EBV and abrogation of that by PWM strongly suggest that at least the fraction of B cells activatable by EBV is refractory to the action of CSA. We interpret the abolition of PWM responsiveness to mean that the site of action of CSA is the T lymphocyte or, less likely, the monocyte—both essential for a PWM-induced polyclonal immunoglobulin response.

Furthermore, CSA seems to promote the early and complete outgrowth of EBV B-lymphoblastoid cell lines from immune donors *in vitro*. It has been shown that at a cell density of 10^6 ml^{-1}, lymphoblastoid cell line generation from EBV-infected lymphocytes from immune donors is inhibited by the *in vitro* appearance of immune T lymphocytes[7]. It seems that CSA specifically deletes or functionally inactivates these T cells, allowing the rapid appearance of EBV-transformed cell lines from all immune donors tested. In six out of nine seropositive individuals, EBV-bearing cell lines have been directly established from the blood merely by the addition of CSA alone. It is known that EBV-infected B cells persist in the lymphoid system of immune donors and can grow out into lines in appropriate conditions if T lymphocytes are removed[7]. CSA seems to facilitate this outgrowth, probably by its specific action on the T lymphocyte.

The experimental conditions we have used are certainly unfavourable for *de novo* infection of B cells by *in vitro* released infectious particles[10], because cells were grown in human AB anti-VCA positive serum. In the presence of specific antibody EBV cannot activate B lymphocytes[11]. Although we cannot exclude transfer of virus between cells in intimate contact we suggest that the lymphoblastoid cells growing out from CSA-treated blood of EBV immune donors really represent the progeny of the latently infected cell.

There have been recent disturbing reports of lymphomas presenting in renal allograft recipients treated with CSA and there is a general tendency for anti-EBV titres to rise in patients on treatment with CSA. In one reported case lymphoma followed asymptomatic infectious mononucleosis[12] and in another EBNA was present in the lymphomatous B cells[3]. On the basis of these *in vitro* experiments we suggest that in the presence of CSA, T-lymphocyte surveillance of EBV-infected B cells is no longer effective. The result is a polyclonal proliferation of EBV-transformed B cells which *in vitro* from B-lymphoblastoid lines but *in vivo* manifest as the polyclonal immunoblastic proliferation characteristic of immunosuppressed patients[13].

We thank Kicki Hild for technical assistance, Kenneth Nilson for characterizing the lines and G. Klein for supplies of B95-8 supernatant. This work was supported by the Swedish MRC and in part by a grant from the North of England Cancer Research Campaign.

Received 2 September; accepted 7 November 1980.

1. Borel, J. F., Feurer, C., Mognée, C. & Stähelin, H. *Immunology* **32,** 1017–1025 (1977).
2. Calne, R. Y. *et al. Lancet* **ii,** 1033–1036 (1979).
3. Crawford, D. H. *et al. Lancet* **i,** 1355–1356 (1980).
4. Leoni, P., Garcia, R. C. & Allison, A. C. *J. clin. Lab. Immun.* **1,** 67–72 (1978).
5. Janossy, G., Gomez de la Concha, E., Waxdal, M. J. & Platts-Mills, T. *Clin. exp. Immun.* **26,** 108–117 (1976).
6. Bird, A. G. & Britton, S. *Scand. J. Immun.* **9,** 507–510 (1979).
7. Moss, D. J., Rickinson, A. B. & Pope, J. H. *Int. J. Cancer* **22,** 622–668 (1978).
8. Thorley-Lawson, D. A., Chess, L. & Strominger, J. L. *J. exp. Med.* **146,** 495–598 (1977).
9. Reedman, B. M. & Klein, G. *Int. J. Cancer* **11,** 499–520 (1973).
10. Rickinson, A. B., Finerty, S. & Epstein, M. A. *Int. J. Cancer* **19,** 775–782 (1977).
11. Bird, A. G. & Britton, S. *Immun. Rev.* **45,** 541–567 (1979).
12. Nagington, J. & Gray, J. *Lancet* **i,** 536–537 (1980).
13. Purtilo, D. T. *Lancet* **i,** 300–303 (1980).

Inhibition of *P. falciparum* growth in human erythrocytes by monoclonal antibodies

L. H. Perrin, E. Ramirez, P. H. Lambert & P. A. Miescher

WHO Research and Training Centre, Department of Medicine,
Hôpital Cantonal, 1211 Geneva 4, Switzerland

Malaria is increasing in incidence and prevalence in most tropical areas and is a major problem for both individuals and communities[1]. Current malaria research is aimed at developing vaccines[2–4] and, for this, it may be useful to define *Plasmodium* antigen(s) related to the development of a protective immune response in the host. Monoclonal antibodies have recently been shown to interfere with rodent malaria infection (*Plasmodium berghei*) at the sporozoite[5] or merozoite[6] stage. We have now raised monoclonal antibodies against single antigenic determinant(s) of *Plasmodium falciparum* and report that some of them inhibit the growth of erythrocytic forms of *P. falciparum in vitro*.

Red blood cells (RBCs) parasitized with *P. falciparum* were obtained from a European patient on his return from Senegal and were adapted and maintained in *in vitro* culture according to

Table 1 Inhibition of *P. falciparum* growth by monoclonal antibodies

		% Parasitaemia		
Culture medium + 10% ascitic fluid from various hybridomas	Ig class	Total	Intact parasites	% Growth inhibition
Hb28 cl1	IgG3	0.2	0	100
Hb26 cl3	IgG2a	1.5	0.6	96
Hb31 cl2	IgG3	1.2	0.5	95
Hb12 cl4	IgG1	9.6	9.4	19
Hb9 cl1	IgM	9.7	9.5	18
Medium + 5% NMS		9.8	9.6	17
Medium + 10% NaCl		11.8	11.4	0
P3-X63-Ag8 (Fo)		9.6	9.4	19

The inhibition of *P. falciparum* growth *in vitro* by various monoclonal antibodies was analysed by a modification of the technique described by Wilson and Phillips[14]. Synchronized cultures of ring forms of *P. falciparum* were diluted with uninfected human RBCs to provide a final suspension of 1.5% RBCs and 0.5% parasitaemia. Aliquots (150 μl) of this suspension were placed in 96 flat-bottomed wells (3040 Falcon plates). Each microculture was done in triplicate. The medium used for the culture (RPMI 1640 with 10% normal human serum)[7] was supplemented with either 10% ascitic fluid produced by various hybridomas or as control with 5% serum of normal BALB/c mice (NMS), 10% isotonic NaCl, or ascitic fluid induced by myeloma cells used for fusion. The culture medium was removed from the settled cells each day and replaced with fresh medium. The plates were incubated at 37 °C in a candle jar[9]. Parasitaemia was determined on Giemsa-stained thin blood smears after 80 h of incubation. In these conditions, more than two cycles of replication took place in the wells containing the medium +10% NaCl. At the end of the experiment, some of the parasites were pyknotic, especially those in cultures supplemented with Hb28 cl1, Hb26 cl3 and Hb31 cl2. Therefore, the percentage inhibition was determined over the number of intact parasites. The results reported here represent three hybridomas with inhibitory activity and two out of a group of eight hybridomas without any significant inhibitory activity (10–19%).

the method of Trager and Jensen[7]. These cultures were not synchronized and contained all the various developmental erythrocytic stages (rings, trophozoites, schizonts, merozoites). Synchronization was induced by two treatments with 5% mannitol at an interval of 34 h (ref. 8); this treatment lysed the mature forms (trophozoites and schizonts), leaving only the ring forms. Approximately 30 h after treatment with mannitol, merozoites were released from schizonts and non-infected RBCs were invaded.

In the first series of experiments, 8-week-old female BALB/c mice were immunized intraperitoneally (i.p.) three times at 2-week intervals with 100 μg of merozoite and schizont antigen in Freund's incomplete adjuvant. Four days before the fusion experiment a booster of 5 μg of the antigenic preparation was administered intravenously (i.v.). Merozoites and schizonts were purified from a synchronized culture of *P. falciparum* by Plasmagel sedimentation[9] and contained < 10% normal RBCs. In a second series of experiments, BALB/c mice were infected i.p. with 1×10^5 erythrocytic forms of *P. berghei*, treated with chloroquine[10] and 3 weeks later injected i.v. with 5×10^5 erythrocytic forms of *P. berghei* 4 days before the fusion procedure. For each experiment, 2×10^7 immune spleen cells in suspension were fused by polyethylene glycol with 1×10^7 non-secreting mouse myeloma cells derived from P3-X63-Ag8 (clone FO given by Dr Fasekas, Basel Institute of Immunology). Conditions for fusion were essentially as described by Galfre *et al.*[11]. Positive hybrids were clone purified twice on mouse macrophage monolayers by limiting dilution and further grown in T75 Falcon flasks. 5×10^6 cells were injected i.p. into BALB/c mice which had been primed 1 week previously with 0.5 ml Pristan i.p.

An indirect immunofluorescence technique was used to assay hybrid supernatants. Acetone-fixed *P. falciparum*-parasitized human RBCs or *P. berghei*-parasitized mouse RBCs were used as antigenic preparation and rabbit fluoresceinated anti-mouse $F(ab')_2$ as the developing second antibody[10,12]. Antibodies specific for the surface antigen or merozoites were similarly detected by reacting hybrid supernatants or ascitic fluid with a suspension of unfixed schizonts and merozoites (~10^5 per 25 μl) purified by Plasmagel[9].

Spleen cells from nine mice immunized with *P. falciparum* merozoites and schizonts and two mice immunized with *P. berghei* were fused in 11 experiments: 47 of 978 supernatants tested reacted with *P. falciparum*-parasitized RBCs; two of these supernatants obtained from mice immunized with *P. berghei* reacted with both *P. berghei* and *P. falciparum*. From each fusion experiment, hybrid cells from one or two wells (total 15) were selected, clone purified twice on macrophage monolayers and grown in large amounts. The cells were injected i.p. into mice, and the ascitic fluid was collected. When the 15 ascitic fluid samples were tested by indirect immunofluorescence against the different developmental stages of acetone-fixed parasites, 10 samples (including Hb28 cl1 and Hb31 cl2) reacted with trophozoites and schizonts only, two (including Hb26 cl3) reacted with schizonts only, and three reacted with all stages. The ascitic fluid samples were positive by indirect fluorescence at dilutions of 1/5,000–1/100,000.

The culture medium used for *P. falciparum* culture[7] was supplemented with 10% of the various ascitic fluids and the mixtures were tested for their ability to inhibit the multiplication of erythrocytic forms of *P. falciparum*. An 80-h assay corresponding to more than two cycles of multiplication in normal culture conditions was used. Three out of 11 ascitic fluids tested had an inhibitory activity, two of them, Hb26 cl3 and Hb31 cl2, were produced with spleen cells from mice immunized against *P. falciparum* antigens, and the third one, Hb28 cl1, with spleen cells from mice infected with *P. berghei* (Table 1). Sera from

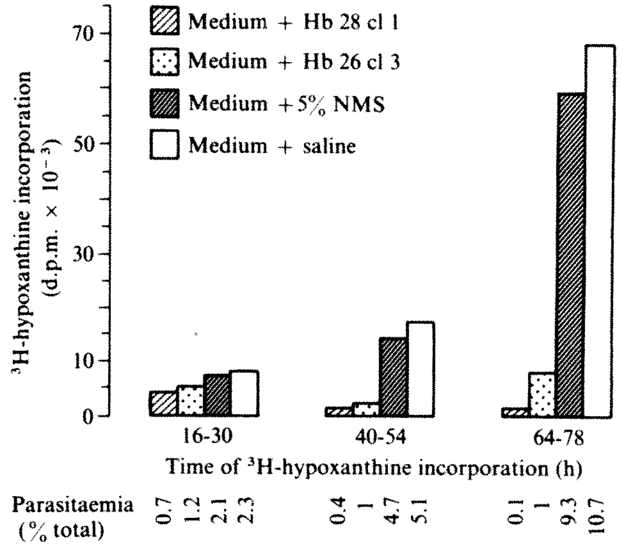


Fig. 1 Incorporation of ^{3}H-hypoxanthine was determined in *P. falciparum* cultures supplemented with 10% ascitic fluid from Hb26 cl3 and Hb28 cl1 hybridomas, 5% BALB/c normal serum and 10% isotonic NaCl, three times during an interval of 14 h. The experimental design was identical to that described in Table 1 legend, except that 10 μl of ^{3}H-hypoxanthine (1 μCi) were added to each culture at the beginning of the labelling period ($18.5–55.5 \times 10^{-3}$ Bq mmol^{-1} specific activity, NEN). The percentage parasitaemia was determined at the end of each labelling experiment on Giemsa-stained thin blood smears.

Table 2 Specificity of the inhibition of *P. falciparum* growth

Culture medium supplemented with:	% Parasitaemia	% Growth inhibition
1. Hybridoma Hb26 cl3		
10% ascitic fluid	0.8	91
IgG-depleted 10% ascitic fluid	7.2	15
Purified IgG fraction (0.5 mg ml^{-1})	0.9	89
2. Mouse immunoglobulins		
IgG2a (LPC-1) (1 mg ml^{-1})	7.8	7
IgG2b (MOPC-195) (1 mg ml^{-1})	9.1	0
Normal mouse IgG (1 mg ml^{-1})	7.7	9
3. Medium + 10% NaCl 0.9%	8.4	0

The specificity of the inhibition of *P. falciparum* growth was analysed in experimental conditions identical to those described in Table 1 legend. The IgG fraction from Hb26 cl3 ascitic fluid was purified by 50% ammonium sulphate precipitation and DEAE cellulose chromatography and shown to be devoid of contaminants when analysed by immunoelectrophoresis. As a control, Hb26 cl3 ascitic fluid was depleted of IgG by passage on a Sepharose–protein A column (Pharmacia) and reconstituted to its original volume by negative pressure dialysis. Mouse immunoglobulin was purified from normal BALB/c sera and the IgG2a and IgG2b from the ascitic fluid of BALB/c mice bearing LPC-1 and MOPC-195 myeloma tumours. These immunoglobulins were prepared by precipitation in 50% ammonium sulphate followed by DEAE column chromatography. The concentration indicated in parentheses for IgG is the final concentration in the medium.

mice injected with hybrid cells of Hb26 cl3 and Hb28 cl1 were also tested and gave, respectively, 78% and 99% inhibition of *P. falciparum* growth when added at a final concentration of 5% to the culture medium.

Hb26 cl3, Hb28 cl1 and Hb31 cl2 monoclonal antibodies inhibited the development of schizonts and the invasion of normal RBCs by merozoites. Normal development of ring forms into trophozoites was observed when differential counts of the developmental stages were monitored on synchronized cultures supplemented with either normal medium or medium containing 10% ascitic fluid from each of the inhibitory monoclonal antibodies. Subsequent maturation of trophozoites into schizonts occurred; however, the schizonts frequently showed morphological signs of degeneration in cultures supplemented with any one of the three monoclonal antibodies. (At the end of the first cycle, these degenerated schizonts represented 15–35% of the expected number of schizonts.) In normal culture, 7.6 new ring forms were generated in the second cycle from each late trophozoite and schizont present in the previous cycle. However, in cultures supplemented with Hb26 cl3, Hb28 cl1 or Hb31 cl2, the number of ring forms per trophozoite and schizont fell to 0.5, 0.2 and 0.4, respectively.

Incorporation of ^{3}H-hypoxanthine during a 14-h pulse was also measured in synchronized cultures supplemented with ascitic fluid. In the absence of ascitic fluid, there was a low rate of incorporation from 16 to 30 h, increased incorporation from 40 to 54 h and the highest rate of incorporation from 64 to 78 h. The increase in incorporation paralleled the increase of parasitaemia. In cultures supplemented with Hb28 cl1 or Hb26 cl3 there was a low rate of incorporation from 16 to 30 h followed by a decrease in incorporation for Hb28 cl1 and a slight increase for Hb26 cl3 from 64 to 78 h. Here again, the rate of incorporation paralleled the level of parasitaemia (Fig. 1).

The IgG fraction of Hb26 cl3 ascitic fluid was shown to have full inhibitory activity (Table 2); the activity of Hb26 cl3 ascitic fluid was lost after depletion of IgG. Similarly, IgG fraction purified from Hb28 cl1 at a final concentration of 0.25 mg ml^{-1} showed an inhibitory activity of 95%. In such cultures inhibited by IgG fractions, schizont degeneration was also present.

The amount of monoclonal antibodies needed to cause 50% inhibition of the culture growth after two multiplication cycles of *P. falciparum* was 8 μg ml^{-1} of culture medium for Hb28 cl1 (1/80 dilution of ascitic fluid) and 16 μg ml^{-1} for Hb26 cl3. *P. falciparum*-infected cultures were also internally labelled with ^{35}S-methionine for 4 h and lysed by 0.5% Nonidet P40 (ref. 13). The lysate was used as antigen in double immunoprecipitation experiments using monoclonal antibodies (Hb26 cl3, Hb28 cl1 and Hb31 cl2)[13]. The washed immune precipitates were analysed on 10% SDS-polyacrylamide gel electrophoresis gels. By autoradiography it was shown that Hb28 cl1 and Hb31 cl2 monoclonal antibodies reacted with a polypeptide of molecular weight (MW) 41,000 whereas Hb26 cl3 reacted with two polypeptides of MWs 96,000 and 36,000. These polypeptides are probably present on the merozoite surface. This was indicated by results of the indirect fluorescence test when an unfixed suspension of *P. falciparum* was reacted with monoclonal antibodies. Intense surface staining was seen using monoclonal antibodies Hb26 cl3, Hb28 cl1 and Hb31 cl2. In contrast, no surface staining was observed with Hb12 cl4 or Hb9 cl1.

Previous investigators have shown that some sera from individuals living in areas endemic for malaria inhibit the growth of *P. falciparum* culture[14–16], but they have not identified the specificity of the 'protective' antibodies. In the present experiment, antigens recognized by monoclonal antibodies inhibiting the growth of *P. falciparum* have been characterized. Interestingly, monoclonal antibodies with inhibitory activity were obtained from the Hb28 cl1 hybrid after fusion of spleen cells from mice infected with *P. berghei*, and reacted by immunofluorescence with *P. berghei* (trophozoites and schizonts), suggesting the presence of a common antigenic determinant on the surface of *P. falciparum* and *P. berghei* merozoites. Several hypotheses can be raised to explain the inhibitory activity of the monoclonal immunoglobulin. First, these antibodies may prevent the merozoite invasion of RBCs by binding to the merozoite surface and interfering directly or indirectly with the parasite ligand–RBC receptor. Second, they may bind to parasite antigens at the surface of infected RBCs at the schizont stage and thus modify the intracellular parasites' metabolism. Third, because schizonts containing RBCs are known to have an increased permeability, one cannot exclude the possibility that some antibody molecules would reach intracellular parasites at the last stage of development.

These monoclonal antibodies provide new tools for the purification of potentially protective antigens of *P. falciparum*, although the large-scale production of *Plasmodium* antigens is still limited by technical problems. An alternative approach is to use DNA recombinant technology to produce defined malarial antigens. In this respect, the protective monoclonal antibodies may be useful for the identification of *Escherichia coli* colonies secreting putative protective antigens[17].

This work was supported by WHO, the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases and the Swiss National Foundation (grant no. 3.890–0.79).

Received 2 May; accepted 21 November 1980.

1. *3rd Annual Report on Special Program for Research and Training in Tropical Diseases* (WHO, Geneva, 1979).
2. Cohen, S., Butcher, G. A. & Mitchell, G. H. *Adv. exp. Med. Biol.* **93,** 89–112 (1977).
3. Reese, R. T., Trager, W., Jensen, J. B., Miller, D. A. & Tantravaki, R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5665–5669 (1978).
4. Siddiqui, W. A. *Science* **197,** 388–389 (1977).
5. Yoshida, N., Nussenzweig, R. S., Potocnjak, P., Nussenzweig, V. & Aikawa, M. *Science* **207,** 71–73 (1980).
6. Freeman, R. R., Trejdosiewicz, A. J. & Cross, G. A. M. *Nature* **284,** 366–368 (1980).
7. Trager, W. & Jensen, J. B. *Nature* **273,** 621–622 (1978); *Science* **193,** 673–675 (1976).
8. Lambros, C. & Vanderberg, J. P. *J. Parasit.* **65,** 418–420 (1979).
9. Reese, R. T., Langreth, S. G. & Trager, W. *Bull. Wld Hlth Org.* Suppl. 1, 53–62 (1979).
10. June, C. H., Contreras, C. E., Perrin, L. H., Lambert, P. H. & Mischer, P. A. *J. Immun.* **121,** 2154–2160 (1979).
11. Galfre, G., Hove, S. C. & Milstein, C. *Nature* **266,** 550–552 (1977).
12. O'Neill, P. & Johnson, G. D. *J. clin. Path.* **23,** 185–189 (1970).
14. Wilson, R. J. M. & Phillips, R. S. *Nature* **263,** 132–134 (1976).
13. Perrin, L. H., Ramirez, E., Liu, E.-H. & Lambert, P. H. *Clin. exp. Immun.* **41,** 91–96 (1980).
15. Cohen, S., McGregor, I. A. & Carrington, S. P. *Nature* **192,** 733–737 (1961).
16. Phillips, R. S., Trigg, P. I., Scott-Finnigan, T. J. & Bartholomew, R. K. *Parasitology* **65,** 525 (1972).
17. Broome, S. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2746–2749 (1978).

Respective roles of laminin and fibronectin in adhesion of human carcinoma and sarcoma cells

Israel Vlodavsky* & Denis Gospodarowicz†

* Department of Radiation and Clinical Oncology, Sharett Institute, Hadassah University Hospital, Jerusalem, Israel
† Cancer Research Institute and the Departments of Medicine and Ophthalmology, University of California, Medical Center, San Francisco, California 94143

In a previous study[1], Ewing's sarcoma cells and colon carcinoma cells[2] failed to attach to plastic and to dishes coated with various collagen types I, III and IV, but adhered rapidly to extracellular matrix (ECM) produced by cultured corneal endothelial cells. However, although carcinoma cells required ECM and did not adhere to fibronectin, the sarcoma cells adhered and flattened almost equally well on either substrate[1]. Thus, different adhesive proteins may mediate the attachment of sarcoma- and carcinoma-derived cells to extracellular matrices, the most likely being fibronectin[3,4] and laminin[5,6]. Fibronectin stimulates the adhesion of fibroblasts, but not epidermal cells, to collagen type IV (ref. 7) and could mediate the attachment of sarcoma cells. Laminin is confined to the lamina lucida region of basement membranes[6] and has been localized to cellular adhesion sites[8]. Studies of the attachment of epidermal cells *in vitro* (V. P. Terranova, personal communication) suggest that it is adhesive for epithelial cells, and so could have the same role for carcinoma cells as that played by fibronectin for sarcoma cells. Cultured vascular endothelial cells secrete fibronectin[9–11] and laminin[12] into both the ECM and the culture medium, and we report here that pre-exposure of plastic dishes to such conditioned medium induces the attachment and flattening of both human colon carcinoma and Ewing's sarcoma cells. While laminin mediates the attachment and spreading of the former fibronectin is responsible for the attachment and flattening of the latter.

Figure 1 shows the time-course of attachment of carcinoma cells to various substrata. Less than 30% of cells seeded on plastic or fibronectin-coated dishes became firmly attached and no flattening was observed, regardless of time in culture (up to 10 days) and whether or not serum (10%) was present. In contrast, 70 and 80% of the seeded cells firmly attached and flattened within the first hour of incubation on dishes pre-exposed to conditioned medium or coated with ECM, respectively (Figs 1, 2). This was observed regardless of whether or not carcinoma cells were seeded in the presence of serum (10%). The best results, in terms of adhesion and percentage of firmly attached and flattened cells, were obtained with medium from endothelial cell monolayers cultured for more than 8 days. Serum-free conditioned medium collected 5–7 days after the cells had reached confluence also promoted obvious cell attachment and flattening. Similar results were obtained with dishes exposed overnight for 2 h (Fig. 2) to the conditioned medium and seeded with either trypsinized single carcinoma cells (Fig. 2*b*, *c*) or cell aggregates (Fig. 2). In the latter case flattening and migration of peripheral cells were observed within 1 h and led to complete dispersion of aggregates 6–8 h after seeding (Fig. 2*b*). When similar tumour cell aggregates were seeded on ECM-coated dishes, flattening and migration were observed within 30 min (Fig. 2*a*). Pre-exposure of dishes to medium conditioned by human skin fibroblasts, vascular smooth muscle cells, or corneal endothelial cells (Fig. 3), or to fresh medium containing 10% serum supplemented or not with fibroblast growth factor (FGF), did not promote cell adhesion, flattening and migration. Studies of the production of laminin by vascular and corneal endothelial cells have shown that depending on the growth stage (sparse actively growing versus confluent resting cultures) the medium conditioned by corneal cells contains 20- to 100-fold less laminin than medium conditioned by vascular endothelial cells.

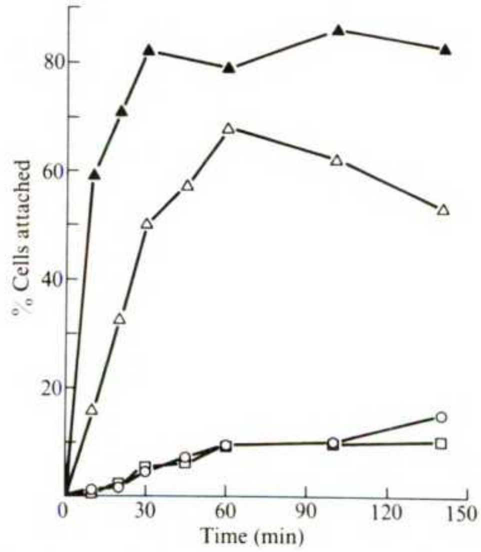


Fig. 1 Time-course of attachment of carcinoma cells to various substrates. Stock cultures of colon carcinoma cells grown on ECM-coated dishes were dissociated by treatment with STV (0.05% trypsin, 0.02% EDTA), and then 2×10^5 cells were seeded on plastic dishes (□), dishes coated with 10 μg each of affinity purified bovine plasma fibronectin (○, ref. 1), dishes coated with ECM (▲, ref. 1), or dishes coated with vascular endothelial cell conditioned medium (△). To coat them with conditioned medium, dishes were exposed to a medium (DMEM, H-16 supplemented with 10% calf serum) taken from vascular endothelial cells maintained for 8 d in culture (from the day of seeding until 3–4 d after reaching confluence). This medium was centrifuged (1,500*g* for 5 min), incubated (14 h, 37 °C) on the dishes, and replaced with fresh conditioned medium before cells were seeded. Serum-free conditioned medium was collected from subconfluent endothelial cultures 7–9 d after they had been supplied with DMEM (H-16) containing no serum. Because the production and secretion of proteins by endothelial cells is greatly reduced when they reach confluence[11,12], medium conditioned by confluent cultures was less active than medium conditioned by actively growing vascular endothelial cells. At various times after seeding unattached carcinoma cells were removed by gentle pipetting and rinsing of the dishes (twice) with phosphate-buffered saline (PBS). The remaining firmly attached cells were dissociated with STV and duplicate cultures were counted with a Coulter counter. The variation in different determinations did not exceed ±10% of the mean.

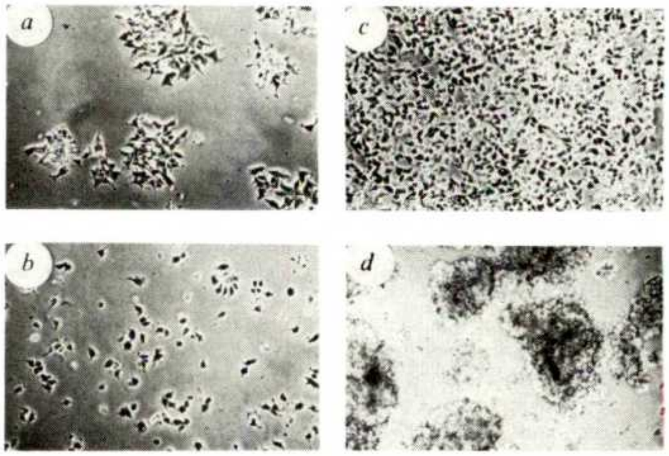


Fig. 2 Morphological appearance of carcinoma cells seeded on dishes pre-exposed to endothelial cell-conditioned medium. Floating cell aggregates (*a*, *d*) from stock cultures maintained on plastic or trypsinized single cells (*b*, *c*) from stock cultures maintained on ECM-coated dishes were seeded in DMEM containing 5% fetal calf serum on plastic (*d*) or on dishes pre-exposed (2 h, 37 °C) to endothelial conditioned medium as described in Fig. 1 legend (*a*, *b*, *c*). *a*, Cell spreading and migration from cell aggregates 8 h after seeding. *b*, Single cells 1 h after seeding. *c*, Confluent culture 6 d after seeding. The cells proliferate and form a layer of firmly attached and flattened cells.

0028-0836/81/040304—03$01.00

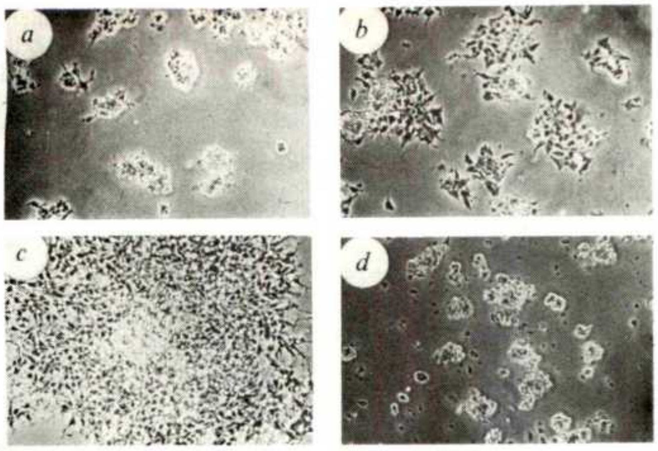


Fig. 3 Effect of conditioned media from various cell types on carcinoma cell attachment and spreading when cells are seeded and maintained in conditioned medium or seeded and maintained in fresh medium. Tissue culture dishes were pre-exposed (14 h, 37 °C) to media conditioned for 8 d by vascular endothelial cells (*a–c*) or corneal endothelial cells (*d*). Carcinoma cell aggregates taken without trypsinization from a culture maintained on plastic were seeded directly into the conditioned medium and observed 6 h later (*a*). Cell aggregates attached loosely to the substrate and only a small degree of cell flattening and migration was observed, even after 1 week in culture. When the conditioned medium was replaced with fresh medium, good cell attachment and spreading were observed after 2 h (*b*). This effect was then long-lasting (*c*, 1 week in culture). This demonstrates that, when carcinoma cells are seeded into conditioned medium-coated dishes, the conditioned medium has to be removed and replaced with fresh medium to promote cell attachment and spreading. Because poor cell attachment and flattening are observed when dishes contain conditioned medium (*a*), it is suggested that, although coating the plastic surface with conditioned medium induces cell flattening, soluble factor(s) which can adversely affect that process is also present in the conditioned medium. *d* shows cell aggregates 4 d after seeding in fresh medium into dishes coated with medium conditioned by corneal endothelial cells. Similar results were obtained with medium conditioned by vascular smooth muscle cells or human skin fibroblasts. In either case, and as was the case with uncoated tissue culture dishes, cell aggregates attached poorly and cell flattening was not observed.

Pre-exposure of dishes to the endothelial cell conditioned medium (either serum-free or containing fibronectin-depleted calf serum) also induced attachment of sarcoma cells. However, in this case and in contrast to carcinoma cells, coating of dishes with purified fibronectin was either more or as effective in promoting cell attachment and subsequent flattening.

The nature of the agents mediating attachment and flattening was analysed by subjecting the conditioned medium to specific double immunoprecipitation with either rabbit anti-laminin[6] or anti-fibronectin antisera[9], followed by goat anti-rabbit IgG antiserum[11]. Experiments with ^{35}S-methionine-labelled vascular endothelial cells had shown that these antisera, as determined by SDS-polyacrylamide gel electrophoresis, precipitate specifically and quantitatively the laminin[12] and fibronectin[11] present in endothelial cell conditioned medium. As Fig. 4*b* shows, carcinoma cells no longer became attached to dishes pre-exposed to conditioned medium if laminin had been removed by immunoprecipitation. On the other hand, immunoprecipitation of fibronectin (Fig. 4*a*) or collagen type IV (not shown) by specific antibodies had no effect on the attachment and flattening of carcinoma cells. In contrast, when the same immunoprecipitated fibronectin-free conditioned medium was used to coat dishes, sarcoma cells no longer became attached (Fig. 4*d*). The induced attachment and flattening of sarcoma cells were not affected by subjecting the conditioned medium to immunoprecipitation with anti-laminin (Fig. 4*e*), anti-collagen type IV or non-immune rabbit serum (Fig. 4*f*) followed by coating the dishes. These results show that the human carcinoma- and sarcoma-derived cells show a specific response to different adhesive glycoproteins, such as laminin and fibronectin, respectively.

Immunofluorescence has shown that the *in vivo* distribution of laminin is identical to that of collagen type IV (refs 6, 8) and so we compared the attachment of carcinoma cells to dishes coated with collagen types I, III, IV or V before being coated with conditioned medium. (Collagen would not promote cell attachment unless dishes were subsequently exposed to conditioned medium.) Collagen type IV gave the best results in terms of adhesion, percentage of firmly attached cells and prevention of retraction of flattened cells within days of initial attachment. Other collagens had either a small effect (type I) or were inhibitory (type II and in particular type V, Fig. 5*d*) compared with direct exposure to conditioned medium. When conditioned medium with a relatively low cell attachment-promoting activity was used (Fig. 5*a*), its effect was potentiated by precoating the dishes with collagen type IV (Fig. 5*b*). These results suggest that laminin may mediate the attachment of human carcinoma cells

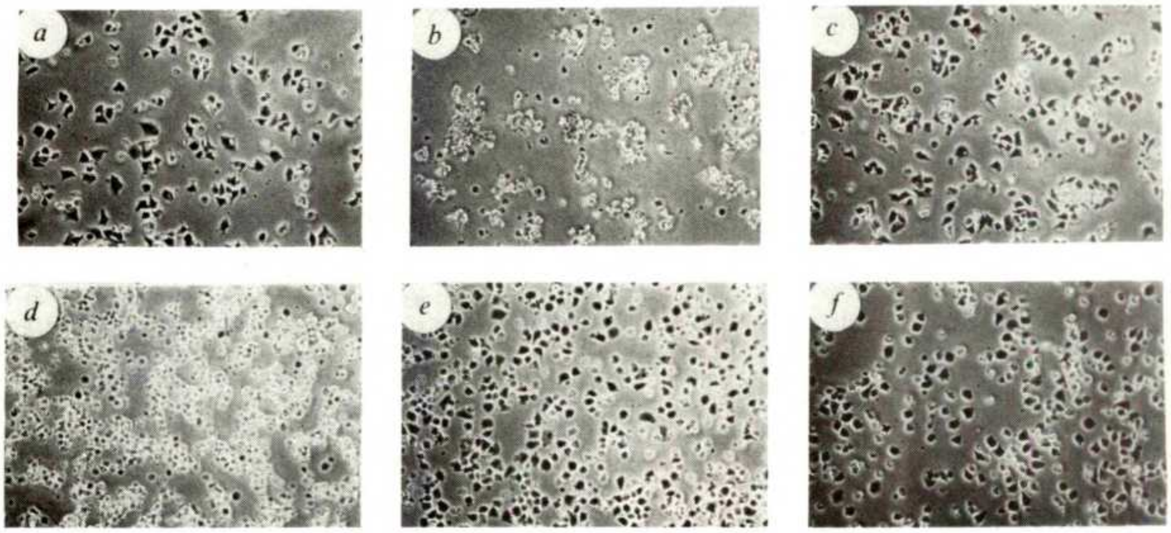


Fig. 4 Induction of tumour cell attachment promoted by coating dishes with vascular endothelial cell conditioned medium from which fibronectin or laminin had been removed by double immunoprecipitation. Vascular endothelial cell conditioned medium (1 ml) was preincubated (1 h, 37 °C) with 10 µl of rabbit antisera against fibronectin (*a, d*) or laminin (*b, e*), or with 10 µl of non-immune rabbit serum (*c, f*). Antibodies to laminin[5,6] and bovine plasma fibronectin[9,11] were the same as those used previously[5,6,9,11]. The specificities of the antisera and lack of cross-reactivity between the various antigens have been established in these previous reports. The conditioned media were then incubated (2 h, 37 °C, followed by 14 h at 4 °C) with 30 µl of goat anti-rabbit IgG antiserum (35% ammonium sulphate cut), centrifuged (12,000*g* for 5 min) and the supernatant incubated in 35-mm tissue culture dishes for 12 h at 37 °C. The conditioned media were then removed and 2 ml of DMEM supplemented with 5% fibronectin-depleted calf serum and containing either 2×10^5 colon carcinoma cells (*a–c*) or Ewing's sarcoma cells (*d–f*) were then added to the dishes. Phase micrographs of cell attachment and spreading were taken 3 h after seeding. Conditioned medium pre-exposed to anti-collagen type IV (refs 5, 6) gave results similar to those observed after pre-incubation with non-immune rabbit serum. Similar results were obtained with serum-free vascular endothelial cell conditioned medium and with tumour cells seeded in DMEM containing 0.1% bovine serum albumin rather than 5% bovine serum.

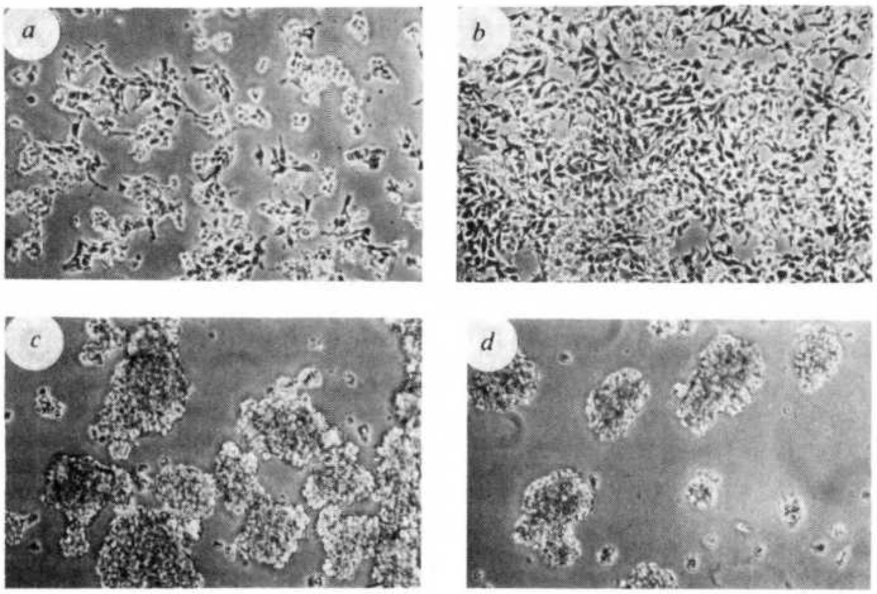


Fig. 5 Effect of precoating with types IV and V collagen on colon carcinoma cell attachment induced by a subsequent exposure to the vascular endothelial cell conditioned medium. Tissue culture dishes (35 mm) were coated with purified preparations of types IV (human placenta) or V (bovine cornea) collagen (10 μg per dish) as previously described[7]. Plastic and collagen-coated dishes were then exposed (12 h, 37 °C) to conditioned medium taken 5 d after replacing the growth medium of subconfluent endothelial cultures. Carcinoma cells (10^5 per dish) were seeded in fresh medium (DMEM, 10% FCS) and phase micrographs taken 3 d later. *a*, Pre-exposure to conditioned medium alone. *b*, Precoating with collagen type IV followed by exposure to conditioned medium. *c*, Dishes coated with collagen type IV alone. *d*, Precoating with collagen type V, followed by exposure to conditioned medium.

to type IV collagen which is characteristic of basement membranes.

Both laminin and fibronectin have been shown to be secreted underneath the endothelial cell layer and to be part of the ECM produced by these cells[12]. In the ECM, they may function as conditioned medium does, and may promote the attachment and rapid spreading of carcinoma- and sarcoma-derived cells[1]. The production of both laminin and fibronectin by vascular endothelial cells may facilitate the arrest of blood-borne metastatic cells[13] that produce little or no such adhesive glycoproteins[14]. Metabolic labelling with ^{35}S-methionine of the Ewing's sarcoma and colon carcinoma cells (cultured on ECM) followed by specific immunoprecipitation had shown that these cells secrete no detectable laminin and small amounts of fibronectin (less than 3% and 0.1% of the newly synthesized proteins, respectively). Interaction of highly metastatic cells with the luminal surface of the vascular endothelium is followed by endothelial retraction and exposure of the underlying ECM, to which the tumour cells adhere strongly[13], possibly through specific attachment-promoting molecules such as laminin and fibronectin. Cell attachment is then followed by migration, degradation and invasion of the tumour cells through the endothelial basement membrane[13,15]. The presence of specific attachment factors in basement membranes, the ability or inability of cells to produce and/or interact with such factors, as well as the preferential interaction of these adhesive factors with a specific type of collagen may determine the localization of a given cell type to specific matrices[7]. Such specific cell–matrix interactions may direct the growth, migration and differentiation of normal cells[16,19], as well as the invasiveness and metastasis of tumour cells[13,15].

The colon carcinoma cell line (HS703T) was provided by Dr R. B. Owens, and the Ewing sarcoma cell line by the laboratory of J. E. Fogh (Sloan Kettering Institute). We thank Dr J. M. Foidart for affinity purified antibodies against laminin and collagen type IV, Dr S. Tseng for collagen types, G. M. Lui for technical assistance, Mr Harvey Scodel for assistance with the manuscript, and Professor Zvi Fuks for support and discussion. This work was supported by a grant from NIH to D.G.

Received 30 June; accepted 6 November 1980.

1. Vlodavsky, I., Lui, G. M. & Gospodarowicz, D. *Cell* **19,** 607–616 (1980).
2. Nelson-Rees, W. A., Hunter, L., Darlington, G. J. & O'Brien, S. J. *Cytogenet. Cell Genet.* (in the press).
3. Klebe, R. J. *Nature* **250,** 248–251 (1974).
4. Yamada, K. M. & Olden, K. *Nature* **275,** 179–184 (1978).
5. Timpl, R., Rohde, H., Robey, P. G., Rennard, S. I., Foidart, J.-M. & Martin, G. R. *J. biol. Chem.* **254,** 9933–9937 (1979).
6. Foidart, J. M. *et al. Lab. Invest.* **42,** 336–342 (1980).
7. Murray, J. C., Stingl, G., Kleinman, H. K., Martin, G. R. & Katz, S. I. *J. Cell Biol.* **80,** 197–202 (1979).
8. Alitalo, K., Kurkinen, M., Vaheri, A., Krieg, T. & Timpl, R. *Cell* **19,** 1053–1062 (1980).
9. Birdwell, C. R., Gospodarowicz, D. & Nicolson, G. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3273–3277 (1978).
10. Jaffe, E. A. & Mosher, D. F. *J. exp. Med.* **147,** 1779–1791 (1978).
11. Vlodavsky, I., Johnson, L. K., Greenburg, G. & Gospodarowicz, D. *J. Cell Biol.* **83,** 468–486 (1979).
12. Gospodarowicz, D., Greenburg, G., Foidart, J.-M. & Savion, N. *J. Cell Physiol.* (in the press).
13. Kramer, R. H. & Nicolson, G. L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5704–5708 (1979).
14. Smith, H. S., Riggs, J. L. & Mosesson, M. W. *Cancer Res.* **39,** 4138–4144 (1979).
15. Liotta, L. A., Tryggrason, K., Garbisa, S., Hart, I. & Folts, C. M. *Nature* **284,** 67–68 (1980).
16. Grobstein, C. *Cancer Inst. Monogr.* **26,** 279–299 (1967).
17. Wessels, N. K. in *Tissue Interactions and Development,* 213–229 (Benjamin, Menlo Park, California, 1977).
18. Gospodarowicz, D., Delgado, D. & Vlodavsky, I. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4094–4098 (1980).
19. Gospodarowicz, D. & Ill, C. R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2726–2730 (1980).

---

cis-Interacting genes in the *S* region of the murine major histocompatibility complex

James Michaelson, Arturo Ferreira & Victor Nussenzweig

Irvington House Institute and Department of Pathology, New York University Medical Center, 550 First Avenue, New York, New York 10016

Insight into the control of gene expression may be gained by analysing genetic systems marked by both regulatory and structural variants. In such systems one can determine whether a regulatory element controls structural genes on both chromosomes or only on the chromosome to which it is linked. The latter may be detected in individuals heterozygous at both the regulatory and structural loci, in which case the effect of each regulatory allele is seen to be exerted only on the *cis*-located structural allele. In prokaryotic organisms, the identification of *cis* interaction of this sort has allowed elucidation of many features of genetic regulation, first for the *lac* operon[1] and subsequently for a variety of other systems[2]. In higher organisms, however, there have been few opportunities to observe *cis*-interacting genes[3–13]. The most thoroughly characterized mammalian system in this regard is the murine β-glucuronidase locus described by Paigen and his colleagues[14–16], in which *cis* interaction has been shown to occur between two closely linked genetic elements—the β-glucuronidase structural gene itself and an androgen-activated regulatory gene which controls the quantity of β-glucuronidase expressed. We report here that *cis*-interacting genetic elements are also found in the *S* region of the mouse major histocompatibility complex *H–2.*

Table 1 *S*-region γ-chain variants detected by IEF

	Haplotype	pI of γ chain
Ss(C4)	$H\text{-}2^{s,p,f,j,w17,w19}$	7.4
	$H\text{-}2^{d,k,r,u,q,b,w7}$	6.9
	$H\text{-}2^{w(\text{undesignated})}$	6.5
Slp	$H\text{-}2^{d,p,u,s,j}$	6.7
	$H\text{-}2^{w7}$	6.5

Results are summarized from refs 25–27.

The *S* region controls two structurally similar serum proteins[17,18], Ss (which has been shown to be C4 of the complement system[19–22]) and Slp (sex-limited protein). Ss(C4) and Slp are each composed of three chains, α (molecular weight (MW) 100,000), β (MW 75,000) and γ (MW 35,000); in both cases the three subunits are derived from 200,000-MW precursors[18,23]. Ss(C4) may be found in two phenotypes—high and low quantity in plasma (a 20-fold difference)—with the *S* region containing the genes determining phenotype[17,24]. We have analysed the Ss(C4) and Slp proteins for biochemical variants which would serve as markers of the structural genes themselves. Recently, we described such variants (herein called *pI* variants)[25–27], identified by the isoelectric point of the γ chain (Table 1). The genetic determination of isoelectric point was found to map to the *S* region, where the previously described regulatory genes are located. Although the most straightforward explanation of the *pI* polymorphism is that of allelic variants of the structural genes, this has not been proved directly and other explanations are possible. Peptide mapping experiments to resolve this question are now underway. In any event, this need not hinder the formal argument of *cis* interaction between *pI* genes and the genes controlling expression.

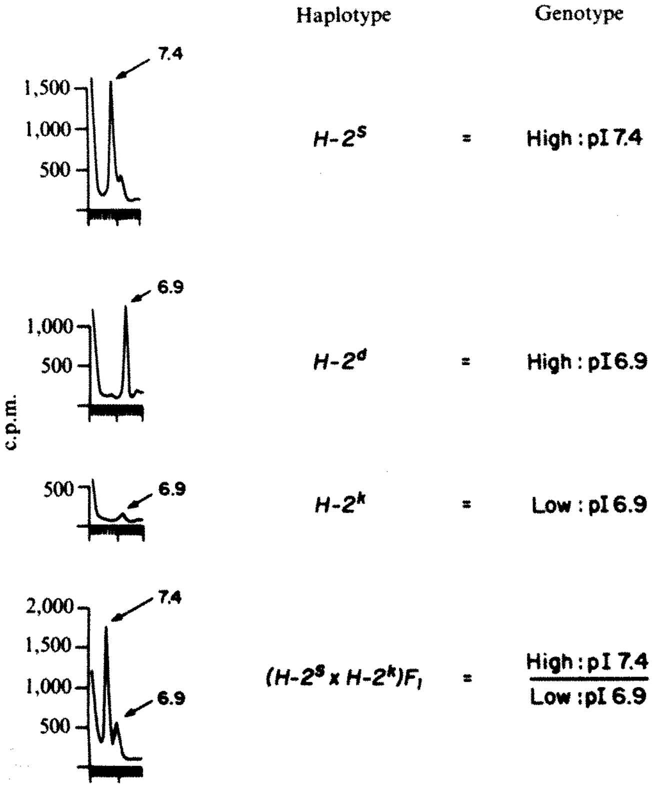


Fig. 1 Cylindrical IEF gels of the γ chain of Ss protein from $H\text{-}2^s$, $H\text{-}2^d$, $H\text{-}2^k$ and $(H\text{-}2^s \times H\text{-}2^k)F_1$ plasma. Ss precipitation was performed after removing Slp from plasma. Analogous results were obtained when Ss was analysed by two-dimensional separation (slab IEF, followed by SDS-polyacrylamide gel electrophoresis).

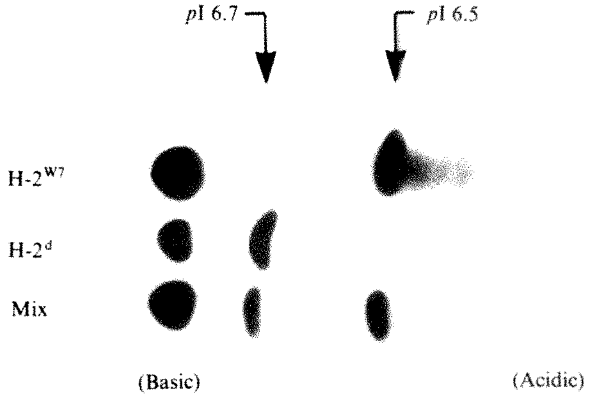


Fig. 2 Slab IEF (*p*H gradient 7.0–3.7) showing Slp γ chain of $H\text{-}2^{w7}$ (C3H.W7) and $H\text{-}2^d$ (DBA/2) mice, and a mixture of the two samples. The identification of the *p*I 6.7 and 6.5 spots as γ chains of MW 35,000 was made by applying individual strips to the top of an SDS-polyacrylamide gel, thus separating the components according to molecular weight (not shown) (see refs 25–27 for technical details).

In our analysis of Ss(C4) expression, we examined the interaction between two genetic elements: that determining the *p*I of the γ chain ($H\text{-}2^s = p$I 7.4, $H\text{-}2^k = p$I 6.9) and that determining the quantity of Ss(C4) ($H\text{-}2^s$ = high, $H\text{-}2^k$ = low). $(H\text{-}2^s \times H\text{-}2^k)F_1$ mice (*high*: *7.4*/*low*: *6.9*) were constructed. It was reasoned that, if the *regulatory* and *pI* genes interacted in *trans* (that is, through a diffusible control substance), equal quantities of 7.4 and 6.9 would be seen. If the interaction was *cis* only, however, a high quantity of 7.4 would be seen, with a low quantity of 6.9, and indeed this was what was found (Fig. 1).

We realize, however, that even though *cis* interaction is formally demonstrated, this might be due to differential labelling or breakdown of the two types of molecules. That the two genes are not precisely the same is seen by the fact that strains of both *Ss high* : *pI 7.4* and *Ss high* : *pI 6.9* type have been detected (Fig. 1). However, we do not know whether the *S*-region *regulatory* and *pI* genes are separate or overlapping entities (in terms of genetic fine structure).

We then performed the same type of analysis for the control of Slp expression. In most Slp$^+$ strains of mice (such as the $H\text{-}2^d$ strain described here), Slp is found in the sera of males but not females[28]. This sex limited expression occurs because androgens are necessary to induce the Slp protein, as has been shown by Passmore and Shreffler[29] (castrated males are Slp$^-$, whereas females which are injected with testosterone propionate become Slp$^+$ (ref. 29)). Mice of an exceptional haplotype ($H\text{-}2^{w7}$), however, are Slp$^+$ in both males and females because expression is independent of androgen, as has been demonstrated by Hansen and Shreffler (such mice continue to produce Slp after castration, and indeed after being made genetically insensitive to testosterone by the *Tfm* mutation[30]). Taken together, these observations identify a hormone-driven regulatory gene controlling the expression of Slp. It has two allelic forms: androgen dependent (*AD*) (that is, $H\text{-}2^d$) and androgen independent (*AI*) ($H\text{-}2^{w7}$). We have found that these two haplotypes also differ at their *pI* genes ($H\text{-}2^d = p$I 6.7, $H\text{-}2^{w7} = p$I 6.5)[27] (Fig. 2). We constructed $(H\text{-}2^d \times H\text{-}2^{w7})F_1$ mice (*AD*: *6.7*/*AI*: *6.5*). If the interaction between the *regulatory* and the *pI* genes occurs by a *trans* mechanism, both males and females would be positive for both *p*I 6.7 and 6.5 Slp's. If, however, interaction occurred in a strictly *cis* fashion, males would be positive for the *p*I 6.7 and 6.5 forms, but females (lacking androgens to turn on *cis*-located *p*I 6.7 product) would express only *p*I 6.5 Slp. Comparison of the isoelectric focusing (IEF) gels of Slp from males and females suggests that this is indeed the case (Fig. 3). Thus, for both Ss(C4) and Slp, the interaction of *regulatory* and *pI* alleles occurs only when they are located *cis*.

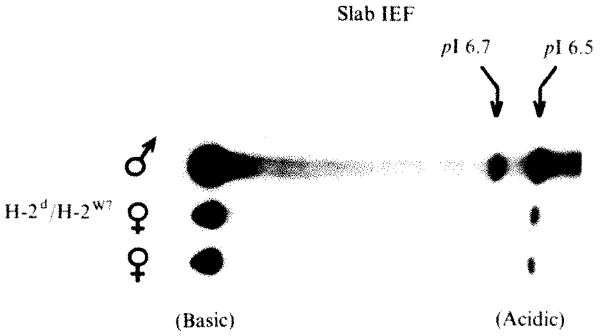


Fig. 3 Phenotype of the γ-chain Slp of (C3H.W7 × DBA/2)F_1 ($H\text{-}2^d/H\text{-}2^{w7}$) = (AD: pI 6.7)/(AI: pI 6.5), seen on slab IEF gels (pH gradient 7.7–4.0). (Components other than the γ chains are nonspecific contaminants.)

As seen in these experiments, the use of IEF to detect *S*-region structural polymorphism permits the simultaneous measurement of both allelic gene products in heterozygotes. From such data, inferences can be drawn about the properties of the genes which control the expression of the variants, as, for example, *cis* restriction of *S*-region regulatory genes. Further experiments of this sort should reveal in greater detail the properties of these regulatory genes. Two features of *S* region expression are noteworthy in this regard. First, because androgens control the serum levels of Ss(C4) (quantitatively)[24] and Slp (absolutely), the analysis of these two proteins may well elucidate some features of the nature of hormone regulation of gene activity. Second, Ss(C4) and Slp have been shown to be synthesized in macrophages and liver parenchymal cells, but not in a variety of other cell types[31]. Presumably, there are genetic elements which control this pattern of synthesis, but where they are or how they exert their effects is unknown. These regulatory elements would be identifiable by tissue-specific genetic polymorphism of Ss(C4) and Slp, if indeed such variation exists. The possibility of such polymorphism is suggested by the observation of Roos *et al.*[18] that *in vitro* synthesis of Slp is seen in macrophages of $H\text{-}2^{w7}$ but not $H\text{-}2^s$ or $H\text{-}2^d$ genotype, although mice of all three types have the protein in their sera. The detailed analysis of this sort of polymorphism, including the use of IEF to test for *cis* and *trans* interaction with *pI* genes, should be informative both for the specific case of the *S* region and with respect to the general question of the control of gene expression during differentiation.

We thank Drs E. A. Boyse, L. Silver and D. Bennett for critical review of the manuscript, and Dr Yuri Bushkin for help with the IEF technique. We also thank Ms Holly Yudkowitz, Mr Victor La Regina and Mr Daniel Eichinger. This work was supported by NIH grants AI-13224, AI-00270 and CA-09149.

Received 23 June; accepted 16 October 1980.

1. Jacob, F. & Monod, J. *J. molec. Biol.* **3,** 318–356 (1961).
2. Miller, E. (ed.) *The Operon* (Cold Spring Harbor Laboratory, New York, 1979).
3. Phillips, R. J. *Genetics* **51,** 485–495 (1966).
4. Glekeler, W., Faversham, J. & Cohn, M. *J. exp. Med.* **148,** 1122 (1978).
5. Dickinson, W. J. & Carson, H. L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4559–4562 (1979).
6. Dickinson, W. J. *Devl Biol.* **42,** 131–140 (1975).
7. Boubelik, M., Lengerova, A., Bailey, D. W. & Matousek, A. *Devl Biol.* **47,** 206–214 (1975).
8. McCaron, M. *et al. Genetics* **91,** 275–293 (1979).
9. Daniel, W. L. *Biochem. Genet.* **14,** 1003–1018 (1976).
10. Carlson, P. *Genet. Res.* **17,** 53–81 (1971).
11. Jarry, B. & Faulk, D. *Molec. gen. Genet.* **135,** 113–122 (1974).
12. Rawls, J. W. & Fristrom, F. W. *Nature* **255,** 738–740 (1975).
13. Lewis, E. B. *Nature* **276,** 565–570 (1978).
14. Paigen, K., Meisler, M., Felton, J. & Chapman, V. *Cell* **9,** 533–539 (1976).
15. Swank, R. T. *et al. Recent Prog. Horm. Res.* **34,** 401–438 (1978).
16. Paigen, K., Labarca, C. & Watson, G. *Science* **203,** 554–558 (1979).
17. Shreffler, D. C. & Passmore, H. C. *Genetics* **48,** 9 (1963).
18. Roos, M., Atkinson, J. P. & Shreffler, D. C. *J. Immun.* **121,** 1106–1115 (1979).
19. Lachmann, P. J., Grennan, D. & Martin, A. *Nature* **258,** 242–243 (1975).
20. Meo, J., Krasteff, T. & Shreffler, D. C. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4536–4560 (1975).
21. Ferreira, A., Nussenzweig, V. & Gigli, I. *J. exp. Med.* **148,** 1186–1197 (1978).
22. Curman, B. *et al. Nature* **258,** 243–245 (1975).
23. Hall, R. E. & Colten, H. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 1707–1710 (1977).
24. Hansen, T. H., Krasteff, T. N. & Shreffler, D. C. *Biochem. Genet.* **12,** 281–293 (1974).
25. Ferreira, A., Michaelson, J. & Nussenzweig, V. *Z. Immun.* (in the press).
26. Ferreira, A., Michaelson, J. & Nussenzweig, V. *J. Immun.* **125,** 1178–1182 (1980).
27. Ferreira, A., Michaelson, J. & Nussenzweig, V. *Immunogenetics* (in the press).
28. Passmore, H. C. & Shreffler, D. C. *Biochem. Genet.* **4,** 351–365 (1970).
29. Passmore, H. C. & Shreffler, D. C. *Biochem. Genet.* **5,** 201–209 (1971).
30. Hansen, T. H. & Shreffler, D. C. *J. Immun.* **117,** 1507–1513 (1976).
31. Saunders, D. & Edidin, M. *J. Immun.* **112,** 2210–2218 (1974).

Detection of CML determinants associated with *H-2* controlled E_β and E_α chains

Antonio Juretić*†, Zoltan A. Nagy* & Jan Klein*

* Max-Planck-Institut für Biologie, Abteilung Immungenetik, Tübingen, FRG
† Department of Physiology, Faculty of Medicine, University of Zagreb, Zagreb, Yugoslavia

The serologically detectable molecules encoded by the *H-2* complex, the major histocompatibility complex (MHC) of the mouse, fall into two classes—class I and class II (ref. 1). The class I molecules encoded by the *K* and *D* loci have a molecular weight of 44,000 and are noncovalently associated with β_2 microglobulin which is not controlled by the MHC. The class II molecules are of two kinds, A and E, each consisting of two noncovalently associated polypeptide chains, α (M_r ~34,000) and β (M_r ~28,000)[2,3]. Three of the four chains, A_α, A_β and E_β, are controlled by loci in the *I-A* subregion, whereas the locus controlling the E_α chain is located in the *I-E* subregion of the *H-2* complex. Thus the loci coding for the E_α and E_β chains are separated by at least one (*J*) and perhaps more loci[3,4]. It has been shown that the E_α and E_β chains are synthesized independently, and that the E_α chain is required for the insertion of the E_β chain in the plasma membrane[4]. We demonstrate here that the $E_\alpha E_\beta$ complex (the E molecule) can evoke *in vitro* cell-mediated lymphocytotoxicity (CML) without previous sensitization *in vivo*, and that the strong CML reactivity is directed against allele-specific determinants on the highly polymorphic E_β chain.

To generate CML against the E molecules, we used three strain combinations selected from the *H-2* recombinant pool such that the responders differed genetically from the stimulators only in the region between the *A* and *S* loci (Table 1). The region may contain, in addition to the loci controlling E_α and E_β chains, the *B*, *J*, *C* and perhaps other, as yet unidentified loci. Four lines of evidence support the conclusion that the CML is directed against E rather than against other molecules encoded by the A–S interval. (1) When tested on a panel of *H-2* recombinant strains, reactivity of some of the effector cells with some of the strains excludes the *B*, *J* and *C* loci from coding for the target-cell determinants. (2) The same panel test indicates that the strain distribution pattern of the determinants detected by the effector cells correlates with the distribution pattern of alleles at E_α- and E_β-encoding loci. (3) The CML typing data fit well with what is known about the expression of E_α- and E_β-encoding loci[3,4]. (4) Cytotoxicity generated across disparities from the *I-B* to *S* region is specifically blocked by a monoclonal antibody against the Ia.7 determinant on the E_α chain[5].

The CML was generated by incubating unprimed spleen cells of the responder strain in a standard mixed-lymphocyte culture (MLC) with γ-irradiated spleen cells of the stimulator strain for 5 days. After incubation, the effector cells were tested on a panel of lipopolysaccharide (LPS)-stimulated, ^{51}Cr-labelled target cells and specific release of radioactivity was taken as a measure of cell lysis. In all three strain combinations, reproducibly high

0028-0836/81/040308—03$01.00

Table 1 Panel testing of primary CML generated against $E_\beta E_\alpha$ molecules

	Target cell *H–2* Regions					% Net ^{51}Cr release by effectors		
Haplotypes	*K*	*I-A* ($A_\alpha A_\beta$)	(E_β)	*I-E* (E_α)	*D*	B10·S(7R) αB10·S(9R) (anti-E^s_β7)	B10·GDαB10·D2(R101) (anti-E^d_β7)	B10·A(4R) αB10·A(2R) (anti-E^k_β7)
k	*k*	*k*	*k*	*7*	*k*	19(16–21)^3	12	37(25–45)^3
a, a1	*k*	*k*	*k*	*7*	*d*	19(12–22)^3	12	41(34–52)^3
h1, h2, h3	*k*	*k*	*k*	*7*	*b*	16(7–23)^9	8(4–10)^4	38(28–47)^{15}
m, m1	*k*	*k*	*k*	*7*	*q*	18(16–19)^2	10	39
aq1	*k*	*k*	*k*	*7*	*f*	17(10–28)^3	11	36(27–44)^2
y1	*q*	*k*	*k*	*7*	*d*	19(7–25)^4	8	40(34–44)^3
by1	*q*	*k*	*k*	*7*	*b*	17(15–20)^3	NT	30(27–33)^2
t1	*s*	*k*	*k*	*7*	*d*	14(6–22)^5	9(6–12)^2	33(24–43)^3
at4	*s*	*k*	*k*	*7*	*b*	15	NT	39(38–40)^2
an1	*s*	*k*	*k*	*7*	*f*	14(8–17)^3	NT	40(35–45)^2
as1	*k*	*k*	*(k)*	*0*	*s*	0(−3–6)^5	4(3–5)^2	3(−2–7)^4
h4	*k*	*k*	*(k)*	0	*b*	0(−7–7)^5	3(0–6)^2	0^7
i3, i5	*b*	*b*	*b*	*7*	*d*	16(10–20)^6	8(6–11)^3	31(21–44)^6
i7	*b*	*b*	*(b)*	*0*	*d*	2(−3–6)^2	4	11(7–14)^3
b	*b*	*b*	*(b)*	*0*	*b*	0(−4–5)^3	1(−2–3)^2	8(5–11)^4
g1, g3	*d*	*d*	*d*	*7*	*b*	11(8–14)^2	21(16–26)^6	19(16–22)^2
o1	*d*	*d*	*d*	*7*	*k*	6(5–6)^2	15(12–16)^3	16(15–16)^2
d	*d*	*d*	*d*	*7*	*d*	6(1–13)^4	20(16–24)^2	17(15–18)^2
g2	*d*	*d*	*(d)*	*0*	*b*	−2(−6–2)^3	0^6	2(−3–6)^4
t3, t4	*s*	*s*	*s*	*7*	*d*	33(24–38)^{11}	11(10–12)^2	19(11–34)^{16}
t2	*s*	*s*	*(s)*	*0*	*d*	0^{11}	5(4–5)^3	5(−1–12)^7
s	*s*	*s*	*(s)*	*0*	*s*	1(−3–5)^3	3^2	3(1–6)^3

CML were set up with 10×10^6 responder and 8×10^6 irradiated stimulator spleen cells in 4.5 ml of Alpha minimal essential medium (Gibco) containing 10% fetal calf serum, and 3×10^{-5} M 2-mercaptoethanol. After 5 days in culture, the cells were tested for lysis of ^{51}Cr-labelled LPS-blasts at different effector : target (E : T) ratios. For designation of target-cell haplotypes see ref. 8. *H–2* regions coding for potential CML determinants are shown. Alleles of the E_β locus in brackets mean failure of cell-surface expression of the relevant polypeptides. Only plateau levels of killing, obtained at E : T 100 : 1 to 160 : 1, are shown. At high E : T ratios, the % net release, that is, the difference between % specific killing of various targets and % specific killing of responder haplotype targets, showed little variation between individual experiments, and was characteristic of the CML combination tested. (Calculation of % specific release was as in Fig. 1 legend.) This fact allowed pooling of data from several experiments. The data are expressed as mean and range of % net release obtained in 2–16 experiments. Superscript numbers show the number of individual experiments (not shown where target haplotypes were tested once).

specific lysis was observed. However, in each combination the reaction was unidirectional—reaction was obtained in the direction B10.A(4R) anti-B10.A(2R) (hereafter 4Rα2R), B10.S(7R) anti-B10.S(9R) (7Rα9R) and B10.GD anti-B10.D2(R101) (GDαR101); all attempts to generate effector cells in the opposite direction have failed. This unidirectional response agrees well with present knowledge concerning the expression of E molecules in the plasma membrane. Jones *et al.*[4] demonstrated that *H–2* haplotypes *b*, *q*, *s* and perhaps also *f* failed to produce functional E_α chains and that this failure prevents the E_β chains from being inserted into the plasma membrane. Thus cells of these four haplotypes bear no E molecules on their surface. In the strain combinations used, one of the strains in each pair always expresses E molecules on the cell surface and the other does not, and we obtained CML only when the responder was the nonexpressor and the stimulator the expressor—never in the opposite direction. In the combination 4Rα2R, unidirectional response was previously observed in MLC, and a complicated explanation involving unequal crossing-over and gene deletion was proposed[6]. Such explanations are no longer necessary; the nonexpression of the E molecule in the membrane fully explains the data—the β-encoding gene cannot be deleted in the B10.A(4R) strain because the E_β chain is present in the cytoplasm of this nonexpressor.

When tested on a panel of strains carrying independent and recombinant *H–2* haplotypes (Table 1), the three effector-cell types gave two kinds of reactivity, one that was comparable in strength with that obtained against the stimulator strain, and one that yielded approximately half as much specific ^{51}Cr release as the lysis of stimulator-type targets (Fig. 1). We shall refer to these two types of response as 'strong' and 'intermediate', respectively. The distinction between the strong and intermediate reactions was also confirmed by cold-target inhibition studies (example shown in Fig. 1).

Table 1 also shows the distribution of alleles at the E_α- and E_β-encoding loci among the tested strains. The E_α locus is shown to possess only two alleles, one designated '*7*' because it

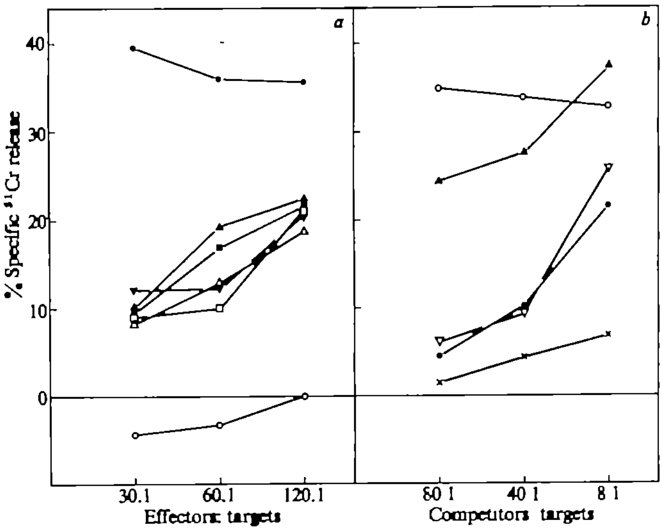


Fig. 1 Reactivity patterns of B10.S(7R) anti-B10.S(9R) effector cells. *a*, Direct cytotoxicity; *b*, cold-target inhibition of cytotoxicity against B10.S(9R) targets. Targets or competitors: ▲, t1 (A.TL); ○, t2 (B10.S(7R)), ▽, t3 (B10.HTT); ●, t4 (B10.S(9R)); △, i3 (B10.A(3R)); □, a(A/J); ■, a(B10.A); ▼, a(A.Tla^b); ×, both targets and competitors are t2. Per cent ^{51}Cr-release was computed using the formula:

$$\frac{\text{c.p.m. supernatant}\times2}{2\times\text{c.p.m. cell lysate}+\text{c.p.m. supernatant}}\times100$$

where the denominator represents total counts per well, when the sample size is half-volume supernatant and half-volume saponin-lysate. Per cent specific ^{51}Cr release was calculated by subtracting per cent release in control groups from that in experimental groups. The controls consisted of labelled target cells mixed with normal spleen cells at ratios corresponding to E : T ratios in the experimental groups. Spontaneous release (control groups) ranged from 17 to 35%, depending on the target strain.

codes for the serologically detectable antigen Ia.7 (ref. 7) and another designated '*0*' because it codes for the absence of Ia.7 in the plasma membrane (however, in reality, there may be only one allele at the E_α-encoding locus and the expression as opposed to nonexpression of the E_α chain in the membrane may be controlled by another locus). The distribution of the *7* allele is that determined by biochemical methods and serological typing for Ia.7 (refs 7, 8).

The E_β-encoding locus is shown to possess at least as many different alleles as there are independent *H-2* haplotypes. This conclusion is based on biochemical data which demonstrate that the E_β chains controlled by independent *H-2* haplotypes differ in their peptide maps[3] (the E_β alleles *b*, *s* and *k* have not yet been distinguished serologically, probably because correct strain combinations for the production of the relevant antibodies have not been used).

The CML combinations 7Rα9R and GDαR101 gave strong CML reactivity only against target cells carrying the *7* allele at the E_α-encoding locus and, at the E_β-encoding locus, the same allele as the stimulator cell (Table 1). As the E_α chains of most Ia.7-positive strains tested derive from H-2^k, the data strongly indicate that these two types of effector cell recognize allele-specific (private) determinants on the E_β chain. The third combination 4Rα2R gave strong CML against target cells carrying the *k* and *b*, but not the *d* and *s* alleles at the E_β-encoding locus, together with the *7* allele at the E_α-encoding locus. The panel test in Table 1 also demonstrates that recognition of the E_β determinants is not restricted by class I antigens, as target cells that share the E molecule with the stimulator cells, but differ from them at the *K* and *D* region, are lysed with equal efficiency. We conclude that the allelic forms of the E_β chain are not only chemically distinct, but also differ functionally, in terms of recognition by cytotoxic T lymphocytes.

The intermediate CML reactivities in all three combinations depend on the presence of E molecules in the target cells' plasma membrane, and thus correlate with the expression of Ia.7 (E_α chain) by the cells (Table 1). Our most recent results suggest that the determinants detected by intermediate CML reactivity are also found on the E_β chain. So far, direct contribution of the E_α chain to the formation of CML determinants could not be demonstrated[5]. Target cells which carry the *0* allele at the E_α-encoding locus, and thus do not express the E molecule on the membrane, are not lysed. The marginal killing of *b* and *i7* target cells by 4Rα2R effectors is an apparent exception; however, the absence of this reactivity on *g2* target cells maps the determinant(s) recognized to the *K* or *I-A* region.

We have also obtained CML reactivity against the E molecule in the strain combination B10.D2(R107) anti-B10.A(3R). This CML combination involves multiple genetic disparities at the *E* and *C* regions, the *Tla* locus and possibly also at other loci to the right of the *S* region[8,9]. This effector-cell type seemed to preferentially recognize target cells of the *i3* and *i5* haplotypes, which express the *b* allele at the E_β-encoding locus together with the *7* allele at the E_α-encoding locus. In addition, they exhibited cross-reactivities independent of the recognition of the E molecule. The genetic basis for these cross-reactions is being investigated.

In conclusion, our results, together with other data[10], demonstrate that the two gene-controlled E molecule serves as a target antigen for recognition by cytolytic T lymphocytes. The CML determinants seem to be localized primarily, if not exclusively, on the highly polymorphic E_β chain. The implications of these findings with regard to the function of the E molecule in immune responses will be discussed in detail elsewhere[11]. But as the E_β and E_α chains also bear serologically detectable determinants[12], participate in Ir-gene control of the immune response and immune suppression[13], and stimulate mixed-lymphocyte reaction[14], there is obviously no need to postulate separate loci for all these functions.

This work was supported in part by grant Wa 139/10/AI.5 from the Deutsche Forschungsgemeinschaft. We thank Ms Gudrun Labib for technical assistance.

Received 25 July; accepted 21 November 1980.

1. Klein, J. *Science* **203,** 516–521 (1979).
2. Cullen, S. E., Freed, J. H. & Nathenson, S. G. *Transplantn Rev.* **30,** 236–239 (1976).
3. Uhr, J. W., Capra, J. D., Vitetta, E. S. & Cook, R. G. *Science* **206,** 292–297 (1979).
4. Jones, P. P., Murphy, D. B. & McDevitt, H. O. *J. exp. Med.* **148,** 925–939 (1978).
5. Juretić, A. *et al. Immunogenetics* (submitted).
6. Lozner, E. C., Sachs, D. H., Shearer, G. M. & Terry, W. D. *Science* **183,** 757–759 (1974).
7. David, C. S. & Shreffler, D. C. *Transplantation* **18,** 313–321 (1974).
8. Klein, J., Flaherty, L., VandeBerg, J. L. & Shreffler, D. C. *Immunogenetics* **6,** 489–512 (1978).
9. Egorov, J. K., Apt, A. S. & Vedernikov, A. A. *Immunogenetics* **5,** 285–290 (1977).
10. Fischer Lindahl, K. & Hausmann, B. *Immunogenetics* (in the press).
11. Klein, J. *et al. Nature* (submitted).
12. Lafuse, W. P., McCormick, J. E. & David, C. S. *J. exp. Med.* **151,** 709–715 (1980).
13. Benecerraf, B. & Germain, R. N. *Immun. Rev.* **38,** 70–119 (1978).
14. Peck, B. A., Klein, J. & Wigzell, H. J. *J. Immun.* **125,** 1078–1086 (1980).

Nonenzymatic hydroxylations of proline and lysine by reduced oxygen derivatives

Robert L. Trelstad, Karen R. Lawley & Lewis B. Holmes

Experimental Pathology and Embryology Laboratories, Shriners Burns Institute,
Departments of Pathology and Pediatrics, Massachusetts General Hospital, and
Harvard Medical School, Boston, Massachusetts 02114

The biosynthesis of *trans*-3- and *trans*-4-hydroxyprolines and 5-hydroxylysine in animal cells requires polypeptide proline or lysine, enzymes and cofactors including iron, and possibly involves peroxidatic intermediates[1]. Several laboratories have reported the presence of low-molecular-weight hydroxyproline and hydroxylysine peptides in cell and organ cultures[2–5]. We found that these small peptides contained the *trans*-3 and *cis*-4 isomers of hydroxyproline as well as *trans*-4 ones and that their production was not completely inhibited by α,α-dipyridyl, an iron chelator and effective inhibitor of enzyme-mediated hydroxylations[5]. It is known that oxygen or hydrogen peroxide in the presence of metal can hydroxylate proline and other aromatic compounds[6–11]. We show here that reduced oxygen derivatives can hydroxylate both free and polypeptide-bound proline and lysine, and that scavengers of hydroxyl radicals suppress, but do not completely inhibit, this reaction. Reduced oxygen derivatives can be generated in normal and pathological circumstances[12], and some of the low-molecular-weight hydroxyproline and hydroxylysine peptides found in cell and organ cultures might be derived from these derivatives and therefore do not reflect collagen turnover, but rather some other cellular activity.

Free proline and lysine and polypeptide-bound forms of each, as well as three isomers of hydroxyproline, were exposed to several different reaction conditions (listed in Table 1), which contained or could generate superoxide anion, hydrogen peroxide or hydroxyl-free radicals[12].

The reaction products were identified using the following techniques: high-voltage electrophoresis[13], thin-layer chromatography[14], chemical derivatization and partitioning by solvent extraction followed by chromatophore generation (including characterization of the spectral absorption pattern of standards and the derivatives)[15], nitrous acid resistance[16], and behaviour on the amino acid analyser, including the reaction product absorption ratios at 440 and 570 nm and co-chromatography with known standards in at least two different elution conditions (Fig. 1).

The products of the reactions involving proline include at least four different hydroxyproline isomers—*trans*-3-hydroxy-

0028-0836/81/040310—03$01.00

Table 1 Nonenzymatic hydroxylations of proline and lysine in free and polypeptide forms

Substrate	Reaction conditions	Reaction products*				
		trans-3-hyp	*trans*-4-hyp	*cis*-3-hyp	*cis*-4-hyp	5-hylys
Free forms:						
	Standard†	2.7	2.6	4.2	3.2	5.4
L-proline	Standard except at 20 °C	1.9	1.8	2.1	1.5	4.3
	Standard except at 0 °C	1.5	1.3	1.9	1.3	4.1
L-lysine	Standard except Fe^{2+} replaced at same concentration by					
	Zn^{2+}	2.0	2.2	4.6	3.0	1.9
	Cu^{2+}		(*trans*-3 + unidentified)			2.6
	Hg^{2+}, Ca^{2+}, Mn^{2+}, Mg^{2+}, Pb^{2+}	‡	‡	‡	‡	‡
	Standard except proline increased fourfold	2.3	3.1	6.4	3.3	
	Standard except Fe^{2+} increased 10-fold	7.3	15.3	22.8	30.6	31.0
	Standard with the addition of free radical scavengers:					
	50 μM mannitol	2.4	1.9	2.9	1.8	1.1
	50 μM sodium formate	1.9	1.6	2.5	1.5	3.0
	5 μM thiourea	2.2	2.1	3.0	2.2	0.1
	10 μM sodium benzoate	1.9	1.5	2.8	1.6	TR
	50 μM potassium chloride	3.1	1.6	4.1	3.3	2.7
	50 μM dimethyl sulphoxide	1.3	1.1	1.6	0.4	TR
	100 μM dimethyl sulphoxide	0.4	0.4	0.4	0.2	TR
	Standard with the addition of iron chelators:					
	10 μM α,α-dipyridyl	3.8	3.0	5.5	3.3	
	300 μM diethylenetriamine penta-acetic acid	3.4	3.2	5.2	3.4	0.6
	Standard without H_2O_2 and with hypoxanthine/xanthine oxidase:					
	with no other additions	2.3	3.5	4.9	5.2	TR
	with added catalase or	0.08	0.10	0.2	0.3	0.0
	superoxide dismutase	0.05	0.06	0.3	0.3	0.0
cis-4-hypro	Standard		6.0		100.0	
trans-4-hypro	Standard		100.0		TR	
trans-3-hypro	Standard		(? 3,4-dihydroxyprolines)			
Polypeptide forms:						
Pro-Gly	Standard	10.5	42.0	20.8	85.6	
Pro-Gly + HNO_2	Standard	2.0	43.4	TR	95.0	
Gly-Pro	Standard	5.7	139.7	TR	TR	
Gly-Pro-Ala	Standard	2.3	2.5	0.7	1.3	
Lys-Lys	Standard					38.7
Lys-Gly	Standard					21.6
Albumen	Standard	TR	9.6	—	—	27.4

TR, trace. *trans*-3-hyp, *trans*-3-hydroxyproline; *trans*-4-hyp, *trans*-4-hydroxyproline; *cis*-3-hyp, *cis*-3-hydroxyproline; *cis*-4-hyp, *cis*-4-hydroxyproline; 5-hylys, 5-hydroxylysine.

* Data were determined by amino acid analyses without hydrolysis (free forms) or after hydrolysis in 6 M HCl (polypeptide forms). The presence of the metal ions did not interfere with the analyses. The values represent the moles of product formed per mole of proline or lysine (or *cis*-4 or *trans*-4) remaining at the end of the reaction ×100. Overall recovery of proline/lysine was less than 10%. The values represent averages of two to six determinations.

† Standard conditions were: 20 μM substrate, 0.2 mM ascorbate, 10 μM $FeSO_4 \cdot 7H_2O$, 1.0 mM H_2O_2, all in 0.1 M NaH_2PO_4, *p*H 4.4, at 37 °C for 1 h, followed by reduction with 1 mM $NaBH_4$ at *p*H 8.0.

‡ These metals with proline gave no reaction; with lysine, 50% of Fe^{2+}.

proline, *trans*-4-hydroxyproline, *cis*-3-hydroxyproline and *cis*-4-hydroxyproline. No differences were noted when D- or L-proline was used in free form. When free proline was the substrate, the relative amounts of each of these isomers produced were nearly equal in the conditions used for most of the experiments. However, with peptide-linked proline, there were significant changes in the distribution of the isomers produced. For example, free proline yields approximately equal amounts of *trans*-3-hydroxyproline and *trans*-4-hydroxyproline, whereas peptide-bound proline is predominantly converted to *trans*-4-hydroxyproline (Table 1). The proline sequence in peptide linkage also affected its relative reactivity and, interestingly, the configuration which is normally hydroxylated in collagen—Pro-Gly—was approximately five times more hydroxylated on a molar basis of starting substrate concentration than Gly-Pro or Gly-Pro-Ala. The proline or lysine recovered at the end of the reaction was less than 10% of the starting material. In addition to the products listed in Table 1, other products were noted and of particular interest was a substance which eluted from the amino acid analyser 7 min before *trans*-4-hydroxyproline, which we have tentatively identified as 4-ketoproline[17]. This material was separated from *trans*-4-hydroxyproline on the amino acid analyser by adjustment of the *p*H of the first buffer to 2.68 (Fig. 1). Several other products, similar in chromatographic properties on the amino acid analyser to the several isomers of the 3,4-dihydroxyprolines, were detected, but not studied further[17,18].

The reaction products generated with free lysine as substrate included 5-hydroxylysine and a second substance which eluted 10 min later, which has not been identified (Fig. 1). With lysine as substrate in polypeptide form, the major product was 5-hydroxylysine.

We suspect that 3- and 4-ketoproline are intermediates in the reaction. The 3-keto form apparently reduces readily in the presence of ascorbate as *trans*-3-hydroxyproline consistently appeared without the addition of sodium borohydride. The 4-keto form was variably present in the initial reaction conditions, but could be quantitatively eliminated by the addition of the sodium borohydride. The 4-keto intermediate was also noted when free *cis*-4-hydroxyproline was used as substrate.

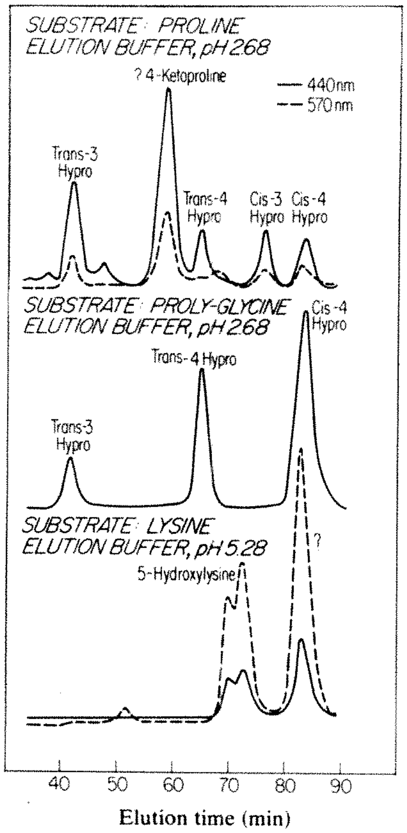


Fig. 1 Elution patterns of reaction products on the amino acid analyser of non-enzymatic hydroxylations of free proline (upper tracing), polypeptide-bound proline which was hydrolysed to free amino acids and then treated with nitrous acid (middle tracing) and free lysine (lower tracing). The solid tracing is the absorbance at 440 nm, the dotted tracing that at 570 nm. The pH 2.68 buffer was 0.18 M in Na^+, the pH 5.28 buffer was 0.35 M in Na^+. The identification of the eluting material by means other than co-chromatography with standards is detailed in the text. The two peaks with ? have not been characterized, but that with proline is probably 4-ketoproline.

When this reaction mixture was reduced, *trans*-4-hydroxyproline was generated. The experiments with hydroxyl radical scavengers, coupled with the fact that no hydroxylations occurred in the absence of added oxygen derivatives, indicate that some form of a reduced oxygen derivative is essential for the hydroxylation reaction to occur. The dependence of the reaction on the presence of metal ions is consistent with a reaction mechanism involving hydroxyl radicals, although the precise nature of the reactants in such mixtures is not understood[19].

Reduced oxygen derivatives have been reported in widely differing circumstances in nature. Of particular interest to matrix biology is the finding that a major effect of these oxygen radicals is to alter extracellular macromolecules. For example, covalent cross-linking via dityrosines in sea urchin fertilization membranes[20], marine mussel byssal fibres[21], collagen[22], elastin[23] and resilin[24] have been reported. Depolymerization of hyaluronic acid has been observed by viscometry following exposure to oxygen radicals[25]. Recently, it has been shown that plasma exposed to superoxide becomes chemotactic for neutrophils[26]; our data suggest that the chemotactic response of fibroblasts and monocytes to hydroxyproline-containing peptides[27] may occur *in vivo* in the absence of collagen breakdown at, for example, sites of inflammation where reduced oxygen derivatives, which we have now shown can hydroxylate non-collagenous peptides, are being produced[12,19,26].

The present data indicate that hydroxyproline- and/or hydroxylysine-containing peptides could be produced *in vivo* by a nonenzymatic reaction involving reduced oxygen derivatives. If this is supported by further studies, then some of the isomers or hydroxyproline which are generated, such as the *cis*-4-hydroxyproline, may be true biosynthetic products. Because *cis*-4-L-hydroxyproline is an analogue of L-proline and effective inhibitor of stable collagen triple-helix formation[28], its production *in vivo* would provide a novel mechanism for regulating net collagen production and add to the growing list of biological effects of oxygen radicals[29].

This work was supported by NIH grants HL 18714, HL 23591, GM 21700 and HD 09689, The March of Dimes Birth Defects Foundation and the Shriners Burns Institute. We thank Mary G. Kirby for technical assistance and Dr Charles Cintron for help with the thin-layer chromatography and high-voltage electrophoresis.

Received 30 April; accepted 16 October 1980.

1. Myllyla, R., Schubotz, L. M., Weser, U. & Kivirikko, K. *Biochem. biophys. Res. Commun.* **89,** 98–102 (1979).
2. Steinberg, J. *J. Cell Sci.* **12,** 217–234 (1973).
3. Nourse, P. N., Nourse, L. D. & Botes, H. *S. Afr. J. Sci.* **70,** 231–234 (1974).
4. Bienkowski, R. S., Cowan, M. J., McDonald, J. A. & Crystal, R. G. *J. biol. Chem.* **253,** 4356–4363 (1978).
5. Holmes, L. B. & Trelstad, R. L. *Devl Biol.* **72,** 41–49 (1979).
6. Chvapil, M. & Hurych, J. *Nature* **184,** 1145 (1959).
7. Lamport, D. T. A. *Nature* **202,** 293–294 (1964).
8. Yip, C. C. *Biochim. biophys. Acta* **92,** 395–396 (1964).
9. Gruber, H. A. & Mellon, E. F. *Analyt. Biochem.* **66,** 78–86 (1975).
10. Udenfriend, S., Clark, C. T., Axelrod, J. & Brodie, B. B. *J. biol. Chem.* **208,** 731–739 (1954).
11. Breslow, R. & Lukens, L. N. *J. biol. Chem.* **235,** 292–296 (1960).
12. Hayaishi, O. & Asada, K. (eds) *Biochemical and Medical Aspects of Active Oxygen* (University Park Press, Baltimore, 1977).
13. Cintron, C., Peczon, B. D. & Kublin, C. L. *Analyt. Biochem.* **87,** 622–630 (1978).
14. Villanueva, V. R. & Lederer, E. *FEBS Lett.* **52,** 308–311 (1975).
15. Kivirikko, K. I. *Int. Rev. Connective Tissue Res.* **5,** 93–163 (1970).
16. Dziewiatkowski, D. D., Hascall, V. C. & Riolo, R. L. *Analyt. Biochem.* **49,** 550–558 (1972).
17. Mauger, A. B. & Witkop, B. *Chem. Rev.* **66,** 47–86 (1966).
18. Irreverre, F. & Witkop, B. *J. Chromat.* **43,** 127–128 (1969).
19. Michelson, A. M., McCord, J. M. & Fridovich, I. (eds) *Superoxide and Superoxide Dismutases* (Academic, London, 1977).
20. Foerder, C. A. & Shapiro, B. M. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4214–4218 (1977).
21. DeVore, D. P. & Gruebel, R. J. *Biochem. biophys. Res. Commun.* **80,** 993–999 (1978).
22. Waykole, P. & Heidemann, E. *Connective Tissue Res.* **4,** 219–222 (1976).
23. LaBella, F., Keeley, F., Vivian, S. & Thornhill, D. *Biochem. biophys. Res. Commun.* **26,** 748–753 (1967).
24. Andersen, S. D. *Biochim. biophys. Acta* **93,** 213–215 (1964).
25. McCord, J. M. *Science* **185,** 529–531 (1974).
26. Petrone, W. F., English, D. K., Wong, K. & McCord, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1159–1163 (1980).
27. Postlethwaite, A. E., Seyer, J. M. & Kang, A. H. *Proc. natn. Acad. Sci. U.S.A.* **75,** 871–875 (1978).
28. Rosenbloom, J. & Prockop, D. J. *J. biol. Chem.* **245,** 3361–3368 (1970).
29. Fridovitch, I. *Science* **201,** 875–880 (1978).

Biosynthetic labelling of platelet activating factor from radioactive acetate by stimulated platelets

H. Chap*, G. Mauco*, M. F. Simon*, J. Benveniste† & L. Douste-Blazy*

* INSERM Unité 101, Hôpital Purpan, 31059 Toulouse, France
† INSERM Unité 200, 32 rue des Carnets, 92140 Clamart Paris, France

Platelet activating factor (PAF-acether) is a chemical mediator released by basophils in *in vitro* and *in vivo* conditions of immunological challenge[1–5]. PAF-acether induces aggregation and secretion of platelets from different species[3,6,7] and could thus be implicated in various immunological forms of tissue injury[8,9]. Other sources have been described for this mediator, including macrophages[10] and platelets[11]. As the amounts of PAF-acether released in the latter case are sufficient to support the platelet aggregation[11] and as platelet aggregation by PAF-acether does not require ADP secretion or thromboxane biosynthesis[12], PAF-acether was proposed as the mediator of the putative third pathway of platelet aggregation[11]. Structural analysis of PAF-acether by lipases indicated that it could be a glycerophospholipid devoid of an ester linkage at the 1-position and bearing a fatty acid esterified at the 2-position[13]. This has provided the first evidence for a phospholipid displaying the properties of a powerful chemical mediator. Based on several

0028-0836/81/040312—03$01.00

lines of evidence, 1-*O*-alkyl-2-acetyl-*sn*-glyceryl-3-phosphorylcholine (1-alkyl-2-acetyl-GPC) was recently proposed as the chemical structure of the mediator, hence the term PAF-acether[14,15]. As PAF-acether synthesis occurs rather early in activated platelets, one of the most efficient pathways of PAF-acether formation could involve a rapid acetylation of the corresponding lysoderivative (1-alkyl-GPC). We now present evidence that platelets activated by ionophore A23187 are able to incorporate acetate into a mixture of 1-alkyl-2-acetyl-GPC and 1-acyl-2-acetyl-GPC; the former supports the biological activity of PAF-acether.

Rabbit blood was drawn on to EDTA (5 mM) by intracardiac puncture and washed platelets were prepared according to Ardlie *et al.*[16]. Final suspension (10^9 platelets ml^{-1}) was in Tyrode's buffer (*p*H 7.35) containing 1 mM $MgCl_2$, 1.25 mM $CaCl_2$ and 0.35% (w/v) fat-free bovine serum albumin (Sigma). Platelets were incubated at 37 °C for 10 min under stirring in the presence of 250 $\mu Ci\ ml^{-1}$ ^3H-acetate (12 Ci $mmol^{-1}$, CEA) and 2.5 μM A23187 (Lilly). Lipids were extracted according to Bligh and Dyer[17]. The lower chloroformic phase was washed three times with methanol/water (1:1 by vol.) to remove excess free acetate. An aliquot of the lipid extract from both control and A23187-treated platelets was subjected to TLC on silicagel F 254 precoated plates (Merck) using chloroform/methanol/water (70:35:7 by vol.) as a solvent. The chromatogram was divided into 17 bands which were directly scraped into scintillation vials containing 10 ml of Picofluor (Packard) and radioactivity was determined by scintillation counting. Figure 1 shows that A23187 induced the appearance of a radioactive lipid migrating very close to sphingomyelin. This chromatographic behaviour is identical to that previously reported for PAF-acether[11].

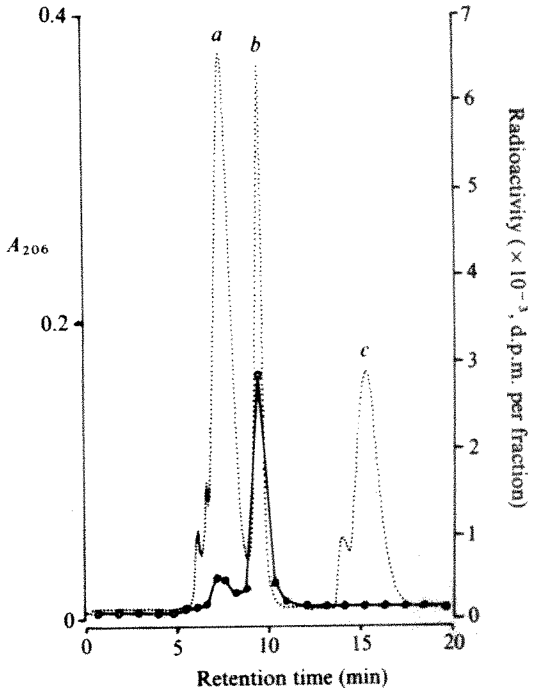


Fig. 2 Identification by HPLC of acetic acid as the hydrosoluble compound released by phospholipase A_2 treatment of purified ^3H-acetate-labelled platelet phospholipid. Preparation of ^3H-acetate-labelled platelet phospholipids has been reported in the text. An aliquot (40,000 d.p.m.) was incubated at 37 °C for 15 min with 5 IU *N. naja* phospholipase A_2 as previously reported[13]. After partition according to Bligh and Dyer[17], to 80 μl of the aqueous upper layer were added 5 μl of pure formic acid (*a*), 5 μl acetic acid (*b*) and 10 μl propionic acid (*c*). The mixture was then subjected to HPLC using a column of RP 18–10 μm (25 × 0.46 cm). Elution was performed according to Farinotti *et al.*[19] at a flow rate of 0.5 ml min^{-1} and 0.5-ml fractions were collected. Carrier acids were detected by absorbance at 206 nm (......) and radioactivity (———) was determined by liquid scintillation counting.

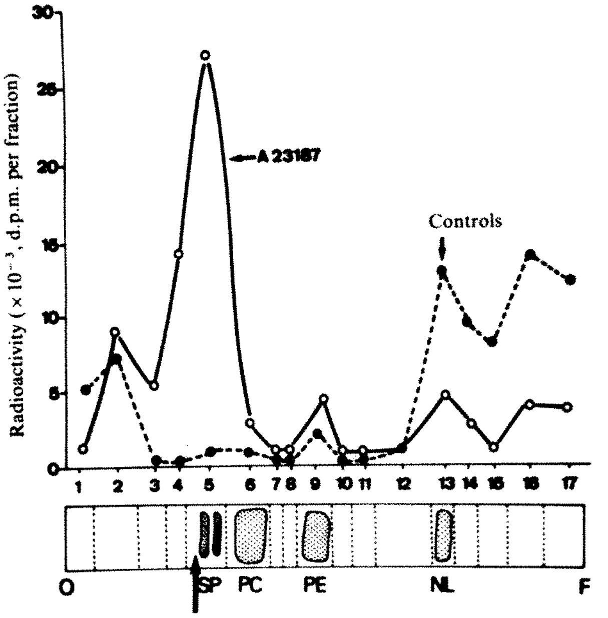


Fig. 1 Distribution of radioactivity after TLC of ^3H-acetate-labelled lipids from platelets. The lipid extracts from control (●) and A23187-treated (○) platelets were chromatographed as described in the text. The lower part of the graph represents the position of major, iodine-stainable lipids and the localization of the different bands scraped into scintillation vials. Each point represents the radioactivity measured in the corresponding band. Results are representative of four identical experiments. O, origin; SP, sphingomyelin; PC, phosphatidylcholine; PE, phosphatidylethanolamine; NL neutral lipids; F, solvent front. The arrow indicates the position of a pure standard of PAF-acether.

To characterize it further, the newly formed radioactive compound was then purified from a similar lipid extract by two steps of HPLC with a Lichrosorb Si 60/5 column (25 cm × 0.9 cm) at a flow rate of 5 ml min^{-1}. Using acetonitrile/methanol/water (60:30:10 by vol.) as a first solvent, a radioactive peak was eluted at 23 min. The pooled fractions (which corresponded to PAF-acether on TLC) were rechromatographed with hexane/propanol 2/water (39:52:9 by vol.). In the latter system, a radioactive peak was detected at 30 min and was clearly separated from other UV-detectable lipids. On TLC, the purified compound exhibited the same R_F as previously observed and it co-migrated with 1-alkyl-2-[^3H]acetyl-GPC and 1-acyl-2-[^3H]acetyl-GPC (see Table 1 legend for their synthesis).

The biological activity of the purified compound was checked on rabbit platelets suspended in Tyrode's buffer (*p*H 7.35) containing gelatin (0.25%, w/v) as previously described[3]. The compound displayed a strong aggregating activity, which still occurred in the presence of creatine phosphate/creatine phosphokinase (4 mM and 10 U ml^{-1}, respectively) or when using platelets incubated for 15 min with 1 mM aspirin before washing (in these conditions, platelets became unresponsive to 10 μM sodium arachidonate).

These results strongly suggest that stimulated platelets synthesize radioactive PAF-acether from labelled acetate. Further biochemical identification was performed by subjecting the purified compound to various treatments. After base-catalysed methanolysis[18] followed by neutralization with HCl and phase partition according to Bligh and Dyer[17], over 90% of the radioactivity was found in the aqueous upper layer. The same result was obtained on incubation with *Naja naja* phospholipase A_2 in previously reported conditions[13]. In these conditions, it was verified by HPLC[19] that the aqueous radioactivity coincided with free acetic acid (Fig. 2). These data confirm that acetic acid was esterified in the 2-position of the isolated phospholipid. However, incubation of the compound

Table 1 Percentage of hydrosoluble radioactivity after lipase treatment of ^{3}H-acetate-labelled phospholipids

	No. of experiments	Controls	Lipase	*P*
Platelet ^{3}H-phospholipid	6	16.2±7.1	60.0±9.6	<0.001
1-acyl-2-[^{3}H]acetyl-GPC	5	16.2±8.0	79.5±4.9	<0.001
1-alkyl-2-[^{3}H]acetyl-GPC	6	29.4±9.7	30.9±8.6	NS

The ^{3}H-acetate-labelled phospholipid isopolar to PAF-acether was purified from a platelet lipid extract as described in the text. 1-Acyl-GPC was obtained by phospholipase A_2 treatment (*Crotalus adamanteus*) of phosphatidylcholine isolated from egg[26] and then purified by TLC. This and synthetic 1-hexadecyl-*rac*-glyceryl-3-phosphorylcholine (Medmark) were acetylated as described by Demopoulos *et al.*[14] in chloroform containing ^{3}H-acetic anhydride (Amersham) diluted with cold anhydride to a specific radioactivity of 1 mCi μmol^{-1}. The two products were then purified by HPLC. Lipase treatment was performed at 37 °C for 60 min in 0.01 M Tris-HCl buffer (*p*H 7.4) containing 0.15 M NaCl and 0.01 M $CaCl_2$. The incubation medium contained in 0.25 ml : 0.02 ml of labelled phospholipid in ethanol (20,000–40,000 d.p.m.) and 0.025 ml of either lipase from *R. arrhizus* (Boehringer) suspended in 3.2 M $(NH_4)_2SO_4$ or 3.2 M $(NH_4)_2SO_4$ in the control tubes. Lipids were then partitioned according to Bligh and Dyer[17] and an aliquot of both lower (organic) and upper (aqueous) phase was determined for radioactivity by liquid scintillation counting using Picofluor (Packard). Results (mean ±1 s.d.) represent the percentage of radioactivity recovered in the upper phase. *P*, probability of significance according to Student's *t*-test. NS, not significant.

with 4 M HCl at room temperature for 3 h (ref. 5) did not modify its phase partition, excluding the occurrence of a vinyl-ether linkage at the 1-position of the phospholipid.

More striking results were obtained following treatment of the radiolabelled compound by lipases. As shown in Table 1, 60% of the radioactivity was water soluble after incubation with lipase from *Rhizopus arrhizus*. This value was close to 80% when incubating 1-acyl-2-[^{3}H]acetyl-GPC, whereas no change in phase partition occurred with 1-hexadecyl-2-[^{3}H]acetyl-GPC. Identical behaviour was obtained when using a lipase with high phospholipase A_1 activity purified from guinea pig pancreas (donated by J. Fauvel and M. J. Bonnefis[20]). These results indicate that the radioactive phospholipid that we isolated from A23187-stimulated platelets could correspond to a mixture of 1-acyl-2-acetyl-GPC and 1-alkyl-2-acetyl-GPC, the former being converted to hydrosoluble 2-acetyl-GPC through lipase hydrolysis. Indeed, TLC results have shown that the water-soluble compound released by lipase treatment of the ^{3}H-labelled platelets was the same as that obtained from 1-acyl-2-[^{3}H]acetyl-GPC. It remained at the origin and was clearly separated from PAF-acether and from free acetate, the latter migrating with the solvent front. In contrast the chromatographic behaviour of 1-alkyl-2-acetyl-GPC was unchanged after lipase treatment, thereby illustrating the specificity of this enzyme towards the ester bond at the 1-position of phosphoglycerides[21]. However, the lower organic phase contained all the original biological activity after such treatment, confirming previous data obtained on PAF-acether from pig leukocytes[13]. We therefore conclude that only 1-alkyl-2-acetyl-GPC supports the PAF-acether activity released by A23187-stimulated platelets. This agrees with the finding that the 1-acyl derivative exhibits a much lower activity than the 1-alkyl derivative[14,15].

The present study is an attempt to characterize the biochemical pathways of PAF-acether synthesis in a living cell. More recent results indicate that a similar incorporation of radiolabelled acetate occurs in stimulated murine macrophages[22]. However, several aspects of acetate incorporation in platelet PAF-acether remain unclear: (1) Are the concentrations of acetate used (20 μM) relevant to available intracellular acetate? Our study does not provide quantitative information; we simply expected the radioactive precursor to label the intracellular pool of acetate, the precise amount of which is unknown. (2) Is the acetylation reaction an enzymatic process involving, for instance, acetylcoenzyme A as an intermediate? This should be the object of further studies. (3) What is the origin of the phospholipid acceptors? 1-alkyl-GPC is released from hog leukocytes[23] and various other cell sources[24], including platelets, but we know of no accurate published report on the putative precursor 1-alkyl-2-acyl-GPC in these cells. The 1-alkyl- and 1-acyl-lysophospholipids could then act as acetyl acceptors for acetate, leading to the formation of a molecule with high or low activity. This would agree with the suppression of PAF-acether synthesis by bromophenacyl bromide, a phospholipase A_2 inhibitor[25]. However, further studies are necessary to characterize the mechanism and the specificity of acetate incorporation into these phospholipids.

This work was supported by INSERM (CRL 78.5.148.3). We thank Dr L. McGregor for correcting the manuscript.

Received 23 July; accepted 31 October 1980.

1. Siraganian, R. P. & Osler, A. G. *J. Immun.* **106,** 1244–1251 (1971).
2. Henson, P. M. & Cochrane, C. J. *J. exp. Med.* **133,** 554–571 (1971).
3. Benveniste, J., Henson, P. M. & Cochrane, C. G. J. *J. exp. Med.* **136,** 1356–1377 (1972).
4. Benveniste, J. *Nature* **249,** 581–583 (1974).
5. Pinckard, R. N., Farr, R. S. & Hanahan, D. J. *J. Immun.* **123,** 1847–1857 (1979).
6. Benveniste, J., Le Couedic, J. P. & Kamoun, P. *Lancet* **i,** 344 (1975).
7. Fesus, L., Csaba, B. & Muszbek, L. *Clin. exp. Immun.* **27,** 512–515 (1977).
8. Henson, P. M. & Pinckard, R. N. *Monogr. Allergy* **12,** 13–26 (1977).
9. Benveniste, J., Camussi, J. & Polonsky, J. *Monogr. Allergy* **12,** 138–142 (1977).
10. Mencia-Huerta, J. M. & Benveniste, J. *Eur. J. Immun.* **9,** 409–415 (1979).
11. Chignard, M., Le Couedic, J. P., Tence, M., Vargaftig, B. B. & Benveniste, J. *Nature* **279,** 799–800 (1979).
12. Cazenave, J. P., Benveniste, J. & Mustard, J. F. *Lab. Invest.* **41,** 275–285 (1979).
13. Benveniste, J., Le Couedic, J. P., Polonsky, J. & Tence, M. *Nature* **269,** 170–171 (1977).
14. Demopoulos, C. A., Pinckard, R. N. & Hanahan, D. J. *J. biol. Chem.* **254,** 9355–9358 (1979).
15. Benveniste, J. *et al. Cr. hebd. Acad. Sci., Paris* **289,** 1037–1040 (1979).
16. Ardlie, N. G., Packham, M. A. & Mustard, J. F. *Br. J. Haemat.* **19,** 7–17 (1970).
17. Bligh, E. G. & Dyer, W. J. *Can. J. Biochem. Physiol.* **37,** 911–918 (1959).
18. Tence, M., Polonsky, J., Le Couedic, J. P. & Benveniste, J. *Biochimie* **62,** 252–259 (1980).
19. Farinotti, R., Caude, M., Mahuzier, G. & Rosset, R. *Analusis* **7,** 449–453 (1979).
20. Fauvel, J. & Bonnefis, M. J. (in preparation).
21. Slotboom, A. J., de Haas, G. H., Bonsen, P. P. M., Burbach-Westerhuis, G. J. & van Deenen, L. L. M. *Chem. Phys. Lipids* **4,** 15–29 (1970).
22. Mencia Huerta, J. M., Roubin, R. & Benveniste, J. (in preparation).
23. Polonsky, J., Tence, M., Varenne, P., Das, B., Lunel, J. & Benveniste, J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
24. Benveniste, J. (in preparation).
25. Mencia-Huerta, J. M., Akerman, C. & Benveniste, J. *Fedn Proc.* **39,** 691 (1980).
26. Papahadjopoulos, D. & Miller, M. *Biochim. biophys. Acta* **135,** 624–638 (1967).

Effects of ATP and vanadate on calcium efflux from barnacle muscle fibres

Mark T. Nelson & Mordecai P. Blaustein

Department of Physiology, University of Maryland, School of Medicine, Baltimore, Maryland 21201
Department of Physiology and Biophysics, Washington University, School of Medicine, St Louis, Missouri 63110

Calcium ions carry the inward current during depolarization of barnacle muscle fibres and are involved in the contraction process[1]. Intracellular ionized calcium ($[Ca^{2+}]_i$) in barnacle muscle, as in other cells, is kept at a very low concentration, against a large electrochemical gradient[2,3]. This large gradient is maintained by Ca^{2+} extrusion mechanisms. When $[Ca^{2+}]_i$ is below the contraction threshold[4], Ca^{2+} efflux from giant barnacle muscle fibres is, largely, both ATP dependent and external Na^+(Na_o^+) dependent (see also refs 5,6). When $[Ca^{2+}]_i$ is raised to the level expected during muscle contraction (2–5 μM)[7], most of the Ca^{2+} efflux from perfused fibres is Na_o dependent; as in squid axons[8], this Na_o^+-dependent Ca^{2+} efflux is ATP independent. Orthovanadate is an inhibitor of ($Na^+ + K^+$) ATPase[9] and the red cell Ca^{2+}-ATPase[10]. We report here that vanadate inhibits ATP-promoted, Na_o^+-dependent Ca^{2+} efflux from barnacle muscle fibres perfused with low $[Ca^{2+}]_i$ (0.2–0.5 μM), but has little effect on the Na_o^+-dependent, ATP-independent Ca^{2+} efflux from fibres with a high $[Ca^{2+}]_i$ (2–

0028–0836/81/040314—03$01.00

5 μM). Nevertheless, ATP depletion or vanadate treatment of high $[Ca^{2+}]_i$ fibres causes an approximately 50-fold increase of Ca^{2+} efflux into Ca^{2+}-containing lithium seawater. These results demonstrate that both vanadate and ATP affect Ca^{2+} extrusion, including the Na_o^+-dependent Ca^{2+} efflux (Na–Ca exchange), in barnacle muscle.

The intracellular solute composition of giant muscle fibres from the barnacle, *Balanus nubilus*, was controlled by internal perfusion[11]. Intracellular $[Ca^{2+}]$ was maintained with Ca-EGTA buffers containing 2 or 8 mM EGTA and appropriate concentrations of $CaCl_2$ (see figure legends). The perfusion fluids also contained caffeine, and the mitochondrial uncoupler, carbonylcyanide *p*-trifluoromethoxyphenyl hydrazone (FCCP), to minimize Ca^{2+} sequestration in sarcoplasmic reticulum and mitochondria, respectively. Caffeine and FCCP do not directly affect Ca^{2+} efflux from internally perfused barnacle muscle fibres (unpublished data). ATP levels are controlled with a regenerating system consisting of phosphoenolpyruvate and phosphokinase. Perfused fibres do not contract even when $[Ca^{2+}]_i$ is raised above the normal contraction threshold (~0.9 μM)[4]. Thus, Ca^{2+} efflux measurements can be made on fibres containing Ca^{2+} levels that would normally be expected to cause contractures (>1 μM)[7].

As in squid axons[8], Na_o^+-dependent Ca^{2+} efflux from barnacle muscle fibres is promoted by ATP when $[Ca^{2+}]_i < 1$ μM (ref. 5). Ca efflux from ATP-depleted fibres perfused with 0.2–0.5 μM $[Ca^{2+}]_i$ was ~0.3 pmol cm^{-2} s^{-1} ($n = 3$) and insensitive to the removal of external Na^+. Similarly, Ca^{2+} efflux from ATP-fuelled, vanadate-treated fibres perfused with 0.2–0.5 μM $[Ca^{2+}]_i$ was ~0.2 pmol cm^{-2} s^{-1} ($n = 3$) and was insensitive to the removal of external Na^+. However, as shown in Fig. 1, in a representative fibre perfused with ATP and 0.2 μM $[Ca^{2+}]_i$, Ca^{2+} efflux was about 0.75 pmol cm^{-2} s^{-1}. This efflux declined to about 0.35 pmol cm^{-2} s^{-1} when external Na^+ was replaced by Tris. When 1 mM vanadate was added to the perfusion fluid, Ca^{2+} efflux into Na^+ seawater fell to ~0.2 pmol cm^{-2} s^{-1}. These data demonstrate that Ca^{2+} efflux from fibres containing low $[Ca^{2+}]_i$ is Na_o^+ dependent, and can be inhibited by vanadate.

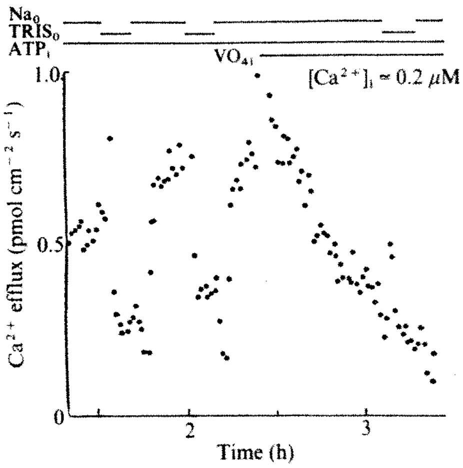


Fig. 1 Effects of external Na^+ and Tris, and internal vanadate on $^{45}Ca^{2+}$ efflux from a fibre perfused with low $[Ca^{2+}]_i$. The composition of the perfusion fluid was: 189 mM K glutamate, 10 mM Na glutamate, 38 mM KCl, 375 mM sucrose, 8 mM $MgCl_2$, 8 mM EGTA, 4.1 mM $CaCl_2$ ($[Ca^{2+}]_i \simeq 0.2$ μM), 3mM Na_2 ATP (Sigma), 3 mM phosphoenolpyruvate (Sigma), 8 μg ml^{-1} pyruvate kinase, and 40 mM HEPES. The perfusion fluids were buffered to *p*H 7.3 with Tris base. A Ca-EGTA stability constant of 7.6×10^6 M^{-1} was used to calculate free calcium concentrations ($[Ca^{2+}]_i$). The composition of the standard external solution (Na_o^+) was (in mmol l^{-1}): 456 NaCl, 6 Tris, 10 KCl, 11 $CaCl_2$ and 32 $MgSO_4$. All external solutions were buffered to *p*H 7.8 at 20°C with maleic acid. Osmolalaties of all internal and external solutions were 930–980 mosmol kg^{-1}. Tris or Li^+ (see Figs 2 and 3) was substituted isosmotically for Na^+ in the Na-free seawaters (indicated by Tris of Lio in the bars at the tops of the figures). Internal perfusion with ^{45}Ca, 2 mM caffeine, and 25 μM FCCP was begun 80 min before the data shown here were obtained. After 150 min of perfusion, 1 mM vanadate was added to the perfusion fluid. Resting potential, −56 mV. Temperature, 16°C. Fibre diameter, 1.4 mm.

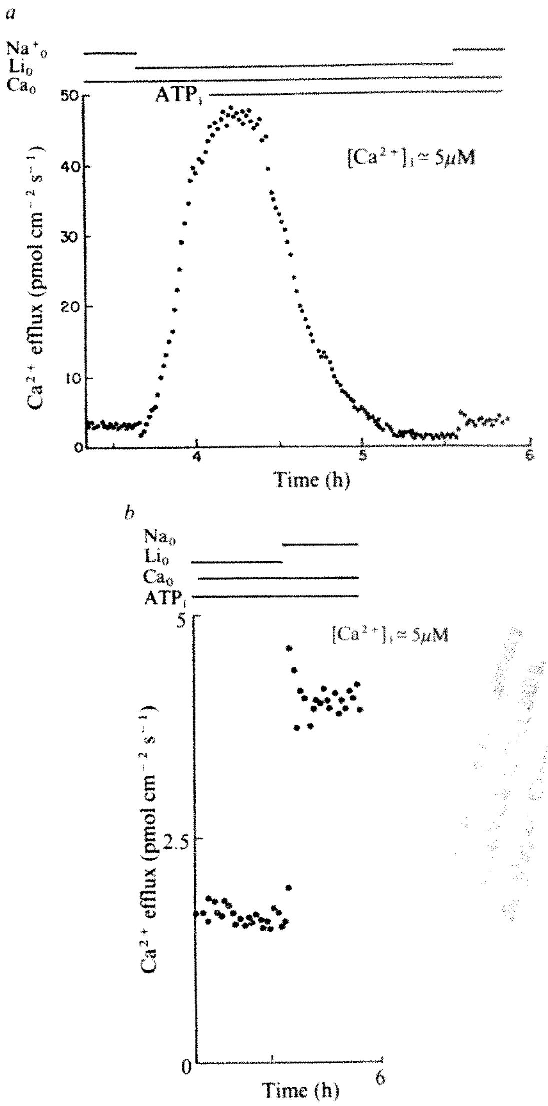


Fig. 2 *a*, Effects of external Na^+ and Li^+, and internal ATP on $^{45}Ca^{2+}$ efflux from a barnacle muscle fibre perfused with high $[Ca^{2+}]_i$. *b*, Effects of external Na^+ and Li^+ on $^{45}Ca^{2+}$ efflux from the same fibre in the presence of ATP (last part of Fig. 2*a* expanded). Internal perfusion with fluid containing $^{45}Ca^{2+}$, 2 mM caffeine, 25 μM FCCP, 2 mM EGTA and 1.95 mM Ca ($[Ca^{2+}]_i \approx 5$ μM) was begun 200 min before data shown here were obtained. After 250 min of perfusion, 2 mM ATP and an ATP-regenerating system were added to the perfusion fluid. The fibre was exposed to 456 mM Na^+ seawater, except during the periods indicated, when the seawater contained 456 mM Li^+. Resting potential, −47 mV. Temperature, 16 °C. Fibre diameter, 1.4 mm.

In fibres containing a high $[Ca^{2+}]_i$, bathed in Ca^{2+}-free media, Li^+ and Tris (and choline) could be used interchangeably as external Na^+ replacements (compare with Fig. 2). However, in fibres with a low $[Ca^{2+}]_i$, Li^+ seems to substitute partially for external Na^+ in supporting Ca^{2+} efflux[12]. Therefore, Tris was used to replace Na^+ in fibres with a low $[Ca^{2+}]_i$ (Fig. 1).

In squid axons containing high $[Ca^{2+}]_i$ (>2 μM), most of the Ca^{2+} efflux is Na_o^+ dependent, but ATP independent[8]. In addition, in high $[Ca^{2+}]_i$, ATP-depleted axons[8], the replacement of external Na^+ by Li^+ increases Ca^{2+} efflux two- to fourfold. This Li_o^+-stimulated efflux is also completely dependent on external Ca^{2+}. Ca^{2+} efflux from perfused barnacle muscle fibres exhibits similar properties. Namely, when ATP-depleted fibres are perfused with high $[Ca^{2+}]_i$, the replacement of external Na^+ by Li^+ increases Ca^{2+} efflux about 10-fold (from ~5 to ~50 pmol

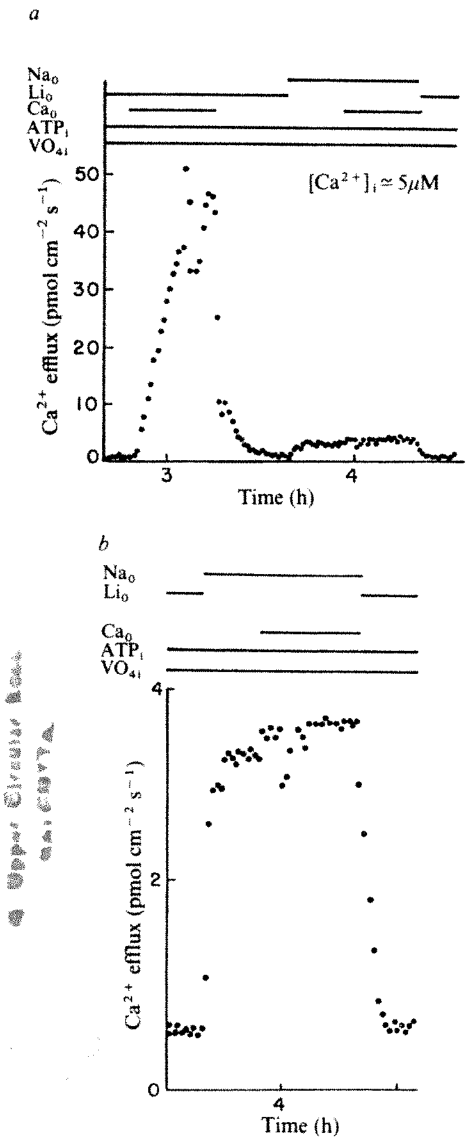


Fig. 3 *a* and *b*, Effects of external Na^+, Li^+, and Ca^{2+} on $^{45}Ca^{2+}$ efflux from an ATP-fuelled fibre perfused with high $[Ca^{2+}]_i$. The end tail of the experiment of *a* is shown on an expanded scale in *b*. Internal perfusion with $^{45}Ca^{2+}$, 3 mM ATP and an ATP-regenerating system, 1.2 mM vanadate, 2 mM caffeine, 8 mM EGTA and 7.8 mM Ca^{2+} ($[Ca^{2+}]_i \simeq 5\ \mu M$) was begun 160 and 210 min before data shown in *a* and *b*, respectively, were obtained. The fibre was bathed in 456 mM Na^+, except for the periods noted, when the seawater contained 456 mM Li^+. The seawater also contained 11 mM Ca^{2+}, except for the periods indicated, when Mg^{2+} replaced Ca^{2+} in the seawater. Resting potential, −45 mV. Temperature, 16 °C. Fibre diameter, 1.5 mm.

cm^{-2} s^{-1}; see Fig. 2*a*). This Li_o^+-stimulated efflux is also entirely Ca_o^{2+} dependent—the removal of Ca_o^{2+} reduces Ca^{2+} efflux into Li^+ seawater (Li SW) to ~1 pmol cm^{-2} s^{-1} (not shown). Moreover, the addition of ATP to the perfusate reduces $Li_o^+ + Ca_o^{2+}$-dependent Ca^{2+} efflux (Fig. 2*a*), and the subsequent replacement of Li_o^+ by Na_o^+ increases Ca^{2+} efflux to about the level seen in the absence of ATP (~4 pmol cm^{-2} s^{-1}; Fig. 2*a*, *b*).

Vanadate mimics the effect of ATP depletion in ATP-fuelled fibres perfused with high $[Ca^{2+}]_i$: it promotes a large Ca^{2+} efflux in Li SW (Fig. 3*a*: compare with Fig. 2*a*). Li_o^{2+}-stimulated Ca^{2+} efflux from vanadate-treated, ATP-fuelled fibres, like Li_o^{2+}-stimulated Ca^{2+} efflux from ATP-depleted fibres, is completely dependent on Ca_o^{2+} (Fig. 3*a*).

Figure 3*a*, *b* also shows that Ca^{2+} efflux from vanadate-treated, high $[Ca^{2+}]_i$ fibres can be promoted by Na_o^+ in the absence of Ca_o^{2+}. In summary, the data from Figs 1 and 3*a* demonstrate that the Na_o^+-dependent Ca^{2+} efflux (presumably Na^+–Ca^{2+} exchange) is blocked by vanadate when it is ATP dependent (in low $[Ca^{2+}]_i$ fibres), but not when it is ATP independent (in high $[Ca^{2+}]_i$ fibres)[11]. Clearly, vanadate cannot be used to distinguish Na_o^+-dependent from Na_o^+-independent Ca^{2+} efflux (contrast with ref. 13).

Resting barnacle muscle fibres have a $[Ca^{2+}]_i$ of about 0.05–0.2 μM (ref. 14). The threshold for contraction is about 0.5–0.9 μM (ref. 4), and maximal isometric force is developed with a 10-fold higher $[Ca^{2+}]_i$ (ref. 7). Our data demonstrate that the Ca^{2+} efflux is Na_o^+-dependent throughout the normal physiological range of $[Ca^{2+}]_i$ values for these fibres. When $[Ca^{2+}]_i$ is below the contraction threshold (< ~0.5 μM)[4], the Ca^{2+} efflux is largely ATP dependent and vanadate sensitive. A substantial fraction of this efflux (~60%) depends on Na_o^+ when $[Ca^{2+}]_i \simeq$ 0.2 μM (Fig. 1). At the $[Ca^{2+}]_i$ levels expected at the peak of contraction (~5 μM)[7], about 90% of the Ca^{2+} efflux is Na_o^+ dependent; but this (Na_o^+-dependent) efflux is neither ATP dependent nor vanadate sensitive.

The mechanism responsible for the residual Ca^{2+} efflux into Na^+-free media is uncertain. This efflux could, in part, be a passive Ca^{2+} and Ca-EGTA leak. However, Ca^{2+} efflux from ATP-fuelled fibres into Na^+-free media was greater than Ca^{2+} efflux from ATP-depleted or vanadate-treated fibres (even with external Na^+ present; compare with Fig. 1). This difference in efflux may represent an ATP-dependent, vanadate-sensitive 'uncoupled' Ca^{2+} pump[13]. Alternatively, this efflux may simply represent an exchange that is activated by other monovalent cations (for example, K^+, Li^+, choline or Tris)[12] and/or by the residual external Na^+ located in the spaces enclosed by surface membrane invaginations.

Our experiments do not show how ATP affects Na^+–Ca^{2+} exchange. However, as vanadate inhibits a number of ATPases[15], our data may infer that the influence of ATP on Na_o^+-dependent Ca^{2+} efflux involves a phosphorylation step. This is consistent with evidence that the non-hydrolysable β–γ methylene analogue of ATP cannot substitute for ATP in promoting Ca^{2+} efflux from the squid axons[16]. Unfortunately, ATP hydrolysis possibly associated with Na^+–Ca^{2+} exchange would be difficult to detect in barnacle muscle fibres which contain many ATPases.

Our data are consistent with the view[8] that ATP increases the affinity of the Na^+–Ca^{2+} exchange system for external Na^+ and/or internal Ca^{2+}. Moreover, the effects of ATP and vanadate on the Ca^{2+} efflux in high $[Ca^{2+}]_i$ fibres may indicate that ATP increases the Na^+/Li^+ selectivity ratio of the carriers, thereby inhibiting $Li_o^+ + Ca_o^{2+}$-dependent Ca^{2+} efflux. Regardless of the uncertainties, our data demonstrate that Na^+ ions, probably acting via an Na^+–Ca^{2+} exchange mechanism, are important for regulating Ca^{2+} in barnacle muscle fibres in physiological conditions.

We thank Dr E. Santiago for technical assistance, Dr L. Reuss for a supply of orthovanadate, and Mrs B. Brooks for preparing the typescript. This research was supported by MDAA, NINCDS, and NSF. M.T.N. was supported by an NIH training grant.

Received 25 August; accepted 30 September 1980.

1. Atwater, I. *et al. J. Physiol. Lond.* **243**, 523–551 (1974).
2. Baker, P. F. *Ann. N.Y. Acad. Sci.* **307**, 250–268 (1978).
3. Blaustein, M. P. *Rev. physiol. biochem. Pharmac.* **70**, 34–82 (1974).
4. Ashley, C. C. *Am. Zool.* **7**, 647–659 (1967).
5. DiPolo, R. & Caputo, C. *Biochim. biophys. Acta* **470**, 389–394 (1977).
6. Russell, J. M. & Blaustein, M. P. *J. gen. Physiol.* **63**, 144–167 (1974).
7. Ashley, C. C. *Ann. N.Y. Acad. Sci.* **307**, 308–329 (1978).
8. Blaustein, M. P. *Biophys. J.* **20**, 79–111 (1977).
9. Cantley, L. C. *et al. J. biol. Chem.* **252**, 7421–7422 (1977).
10. Bond, G. H. & Hudgins, P. M. *Fedn Proc.* **37**, 315 Abstr. (1978).
11. Nelson, M. T. & Blaustein, M. P. *J. gen. Physiol.* **75**, 183–206 (1980).
12. Dipolo, R. *et al. J. gen. Physiol.* **67**, 433–467 (1976).
13. DiPolo, R. *et al. Nature* **281**, 228–229 (1979).
14. Hagiwara, S. & Nakajima, S. *J. gen. Physiol.* **49**, 807–818 (1966).
15. Simons, T. J. B. *Nature* **281**, 337–338 (1979).
16. DiPolo, R. *J. gen. Physiol.* **69**, 795–813 (1977).

Co-release of enkephalin and catecholamines from cultured adrenal chromaffin cells

Bruce G. Livett, Deanne M. Dean, Larry G. Whelan, Sidney Udenfriend* & Jean Rossier*†

Division of Neurology, The Montreal General Hospital and McGill University, Montreal, Quebec, Canada H3G 1A4
* Roche Institute of Molecular Biology, Nutley, New Jersey 07110

The opioid peptides Leu-enkephalin and Met-enkephalin[1,2] are stored intraneuronally in the brain[3] where they are thought to act as neurotransmitters and/or neuromodulators[4]. Evidence for their release from nerve terminals has come from biochemical and pharmacological studies *in vitro* with brain tissue slices[5,6] and synaptosomes[6]. Enkephalins also exist in the peripheral nervous system in nerve cell bodies and axon terminals in the gastrointestinal tract[7], sympathetic ganglia[8–11] and adrenal gland[12]. In the adrenal gland, high levels of enkephalins are present both in axon terminals of the splanchnic nerve[13] and in the adrenal medullary chromaffin cells[13–15] where they are stored together with the catecholamines in the chromaffin granules[14–16]. Stimulation of the adrenal gland *in vivo*[15,17] or the perfused gland *in vitro*[14] causes release of catecholamines and enkephalins into the adrenal vein. However, it is not clear whether the origin of the released enkephalins is the adrenal medullary chromaffin cells or the enkephalin-containing splanchnic nerve terminals that innervate the medulla. We now show that enkephalin and catecholamines are released together from primary cultures of bovine adrenal medullary chromaffin cells by nicotine in a Ca^{2+}-dependent manner.

Primary cultures of bovine adrenal medullary chromaffin cells provide a useful experimental system for investigating the release of catecholamines uncomplicated by diffusion barriers present in intact tissue. For example, this cell culture system has been used to characterize the pharmacological release of catecholamines from the cells by nicotinic agonists[18], to investigate the role of substance P[18,19], somatostatin[19] and enkephalins[20] as modulators of these nicotinic receptors, and to demonstrate the presence of Leu- and Met-enkephalin in high concentrations in the processes and terminal varicosities that characterize these adrenal paraneurones[21]. For the present study, primary cell cultures of adrenal medullary cells were established as monolayers on collagen-coated plastic tissue culture dishes[19] (for further details see Fig. 1 legend). After 8 days in culture, the cells were processed for immunohistochemical localization of Leu-enkephalin as previously described[21]. Preimmune rabbit serum, rabbit anti-Leu-enkephalin serum (R164B) preabsorbed with Met-enkephalin and fluorescein-conjugated goat anti-rabbit serum used in this study were obtained from Dr R. P. Elde. The specificity of the Leu-enkephalin antiserum has been characterized by radioimmunoassay and immunofluorescence[21]. Strong immunofluorescence specific for leu-enkephalin (Fig. 1*b*) was observed in the perikarya and varicose processes of approximately 30% of the adrenal paraneurones whereas over 80% of the cells stained positively for catecholamines (Fig. 1*c*) by the glyoxylic or formaldehyde–glutaraldehyde techniques[18,21]. The immunoreactive Leu-enkephalin had a fine granular disposition within the cytoplasm and processes and exhibited an increasing proximal to distal ratio of fluorescence consistent with synthesis in the perikarya with subsequent transport, processing and storage in the terminals[21].

Catecholamines are normally released from the adrenal gland *in vivo* following activation of acetylcholine (ACh) receptors on the chromaffin cells by ACh released from splanchnic nerve terminals innervating the medulla. In a similar manner, pharmacological studies on bovine adrenal paraneurones have shown that ^{3}H-noradrenaline taken up and stored in reserpine-sensitive sites can be released from the cells by activation of nicotinic (but not muscarinic) ACh receptors[19]. The results in Table 1 show that in the presence of nicotine (ED$_{50}$ 5 × 10^{-6} M) the cells released 13.2 ± 1.6% (n = 5) of their endogenous total catecholamines over a 5-min period. Nicotine produced a preferential release of endogenous noradrenaline (20.9 ± 1.4%) over adrenaline (10.4 ± 0.8%). It also released enkephalins from the cells, and the fraction of the cellular Leu-enkephalin released (8.2 ± 1.3%) was similar to that of the endogenous adrenaline released. The release of both catecholamines and Leu-enkephalin by nicotine was Ca^{2+} dependent (Table 1). Although Met-enkephalin and enkephalin congeners and precursors contained in the chromaffin granules[14–16] are also released[15,22], the combination of immunoassay gel filtration and HPLC methods[22,23] used in our study were highly selective for Leu-enkephalin (see Table 1 legend).

Table 1 Calcium-dependent release of Leu-enkephalin and catecholamines from adrenal paraneurones

	% Cellular content released		
Secretory stimulus	Noradrenaline	Adrenaline	Leu-enkephalin
Nicotine 5 × 10^{-6} M	1.4 ± 0.7	1.0 ± 0.2	ND
Nicotine 5 × 10^{-6} M + Ca^{2+}, 2.2 mM	20.9* ± 1.4	10.4† ± 0.8	8.2*† ± 1.3

The release studies were carried out on 10 culture plates (6 days old) each containing 1 × 10^6 chromaffin cells grown in the presence of Bacitracin (see Fig. 1 legend) in 1 ml medium. Five plates served as controls (nicotine but no Ca^{2+}) and the other five contained nicotine and 2.2 mM Ca^{2+}. Each plate was washed at 37 °C for 15 min in 1.0 ml of a HEPES-buffered Krebs–Ringer saline (KRH) containing 0.2% bovine serum albumen[19] and the plates were then incubated in the release experiments for three successive 5-min intervals in 1 ml of KRH medium alone (0–5 min), KRH medium containing nicotine (5–10 min); KRH medium alone, 'washout' sample (10–15 min). At the end of each 5-min period, the incubation media were removed from the cells, placed on ice and the new incubation medium pre-equilibrated to 37 °C added to the cells. A 100-μl aliquot of the medium removed from the cells was acidified with 25 μl 2 M perchloric acid for catecholamine estimation and the remaining 900 μl acidified with 900 μl 2 M acetic acid for subsequent extraction, chromatography and analysis of Leu-enkephalin. After 15 min the cells were scraped off the plate in 0.5 ml of ice-cold KRH with a rubber policeman and the extract acidified as above for assay of catecholamines and Leu-enkephalin. Assay of endogenous catecholamines: Endogenous catecholamines (both within the cells and released into the media) were assayed by the two-wavelength differential spectrofluorimetric procedure[25] of Renzini *et al*[26]. Noradrenaline and adrenaline levels were obtained by reference to standards after solving simultaneous equations for fluorescence of the samples at 410 and 455 nm with a program (written by Mr Serge La Fontaine) for a digital PDP 11/40 computer. The variability of triplicate assays of the same sample was ±5%. These day-6 cultures contained 7.6 ± 0.8 (n = 10) μg adrenaline per 10^6 cells and 2.9 ± 0.5 (n = 10) μg noradrenaline per 10^6 cells (molar ratio adrenaline : noradrenaline = 2.4). Assay of endogenous Leu-enkephalin: The freeze-dried extracts of cells and media were reconstituted in 10 mM sodium phosphate buffer pH 7.5, 145 mM NaCl supplemented with thiomersal 0.01%, gelatine 0.1% and crystalline bovine serum albumen 0.01%. The extracts were assayed by a radioimmunoassay using a C-terminal-directed Leu-enkephalin antiserum. The characteristics of the C-terminal Leu^5-enkephalin antiserum (RB 92) have been described elsewhere[23]. This antiserum has a cross-reactivity of 3% with Met^5-enkephalin, and is quite specific for the C-terminus of Leu-enkephalin. In terms of molar ratio it has a cross-reactivity of 3% with Met-enkephalin[23] and <1/10,000 with Leu-enkephalin-Arg^6 and Met-enkephalin Arg^6. Note that although the radioimmunoassay is highly selective for Leu-enkephalin, Met-enkephalin is generally present in amounts several times greater than Leu-enkephalin, and is also released[15] and will therefore necessarily be detected in the radioimmunoassay. However, the Leu^5-enkephalin immunoreactivity in the cell extracts and in the media was further characterized by gel filtration and HPLC and was found to correspond to authentic Leu^5-enkephalin. The amount of endogenous catecholamines and Leu-enkephalin contained in the cells at the beginning of each experiment was estimated for each culture plate by summing the amounts in the media at the three time periods and adding it to that remaining in the cells at 15 min. The release of endogenous catecholamines and Leu-enkephalin from the cells into the medium during the 5-min stimulation period is expressed as a percentage of the total content in the cells at the beginning of the experiment. The results give the mean ± s.e.m. (n = 5) where the percentage of catecholamines and Leu-enkephalin released by nicotine stimulation has been compared by Students t-test. ND, not detected.

* $P < 0.001$, † not significant at $P < 0.2$.

A characteristic feature of the nicotinic receptor response is that it shows activation at low levels of agonist (10^{-7}–10^{-5} M nicotine) and inhibition at higher levels (> 10^{-4} M). In accord

---

† Present address: Laboratoire de physiologie nerveuse, CNRS, 91190 Gif sur Yvette, France

0028-0836/81/040317—03$01 00

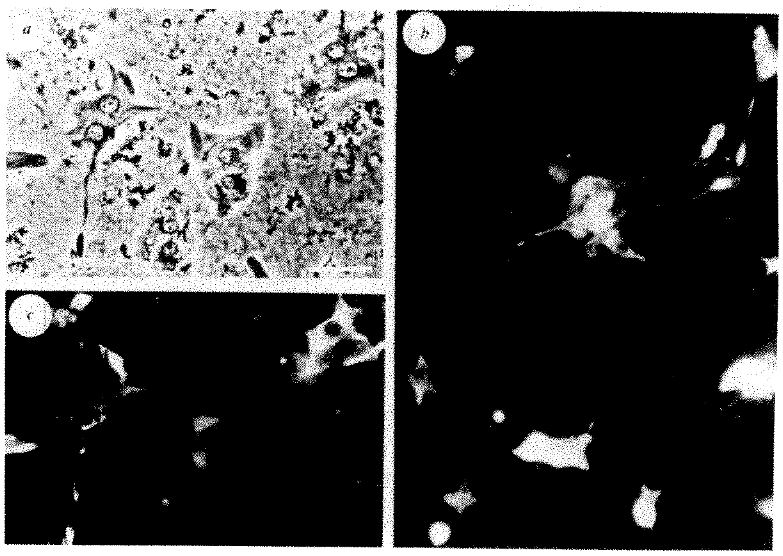


Fig. 1 Histochemical localization of Leu-enkephalin and catecholamines in the perikarya and varicose processes of bovine adrenal paraneurones maintained in cell culture for 6 days, *a*, Phase: background material is serum protein and collagen substrate precipitated by Bouin's fixative. *b*, Same field as *a*; Leu-enkephalin immunofluorescence. *c*, Catecholamine fluorescence. Bars represent 50 μm in *a*, *b*, 100 μm in *c*. Chromaffin cells were isolated from bovine adrenal medullae by retrograde perfusion with Collagenase (Sigma type 1; 0.5 mg ml^{-1})[27] and DNase (Sigma type 1; 13 μg ml^{-1}) in a Ca^{2+}- and Mg^{2+}-free Locke's buffer[28] and purified by fractionation on a self-generating gradient of Percoll (Pharmacia) as described previously[19,29]. By this procedure 195 × 10^6 chromaffin cells of density 1.060.01 g ml^{-1} were obtained from three adrenal glands (wet weight of medullae 25.8 g) with an adrenaline : norrenaline ratio of 3 : 1. The cells were plated in a medium consisting of 80% Eagle's minimal essential medium (MEM) (Gibco), 10% human placental serum (heat-inactivated), 2% 9-day chick embryo extract (50% in balanced salt solution), 600 mg% glucose, cytosine arabinoside (10^{-5} M), 5-fluorodeoxyuridine (10^{-5} M), uridine (10^{-5} M), penicillin (200 mg ml^{-1}) and streptomycin (100 mg ml^{-1})[30], on collagen-coated plastic[31] tissue culture dishes (Corning, 35 mm, 10^6 chromaffin cells) or Aclar 33C plastic 'hats' (Allied Chemical 15 mm, 5 × 10^5 chromaffin cells), and incubated in a humid atmosphere at 36 °C for 6 days before use. The media were replaced at days 2 and 4; and on days 5 and 6 Bacitracin (Sigma) was added to a final concentration of 5 × 10^{-5} M. Immunofluorescence staining of the cells (*b*) with the anti-Leu-enkephalin serum (R164B, 1/90) was shown to be specific by its total inhibition when preabsorbed with synthetic Leu-enkephalin (Peninsula)[21]. Furthermore, cultures incubated with preimmune serum or with the fluorescent conjugate alone showed no specific fluorescence[21]. The cells also contained high concentrations of catecholamine in their perikarya and varicose terminals as demonstrated by the formaldehyde–glutaraldehyde ('Faglu')[32] fluorescence histochemical technique. The final preparations were observed unmounted (for immunofluorescence, in 0.1 M bicarbonate-buffered glycerol; for catecholamine, dry) by incident fluorescence illumination with ×10 or ×20 dry Fluortar (Wild-Leitz) objectives using a Leitz Ortholux-II fluorescence microscope fitted with a Ploem incident illuminator and dichroic mirror primary and secondary filters optimal for fluorescein isothiocyanate or for catecholamine fluorescence. The results were recorded on Tri-X (Kodak) film rated at 400ASA. Exposure times were in the range 5–10 s.

with this general property, the dose–response curve for release of endogenous catecholamines and enkephalins by nicotine (Fig. 2) showed similar responses with inhibition at high concentrations of agonist. At day 6, the adrenal paraneurones contained 699 ± 51 ($n = 7$) pg Leu-enkephalin per 10^6 cells and 20.7 ± 1.7 ($n = 7$) μg catecholamines per 10^6 cells: the molar ratio of catecholamines to Leu-enkephalin released by 5 × 10^{-5} M nicotine (105 × 10^3) was similar to that present in the cells (94 × 10^3).

A recent study[17] showing that Met-enkephalin circulates in human plasma as the intact pentapeptide and that its concentration in the adrenal vein is twice that in the internal jugular or in the peripheral circulation, raises the possibility that enkephalins are secreted together with catecholamines from the adrenal medulla. This is strongly supported by the parallel release of catecholamines and enkephalins recently demonstrated in perfused dog adrenal[15] and by the present study. Further, the close correspondence of release of adrenaline and Leu-enkephalin (Table 1), and the previous immunohistochemical evidence[13] for co-storage of noradrenaline with Met-enkephalin in adrenal medullary gland cells may be of physiological significance. There is increasing evidence that amines and peptides may co-exist not only in adrenal chromaffin cells but also in several nekonal systems[24]. The present demonstration that amines and peptides are released simultaneously from adrenal chromaffin cells in response to a stimulus raises the possibility of co-secretion of catecholamines and enkephalins from adrenergic neurones and the question of the role of this co-secretion.

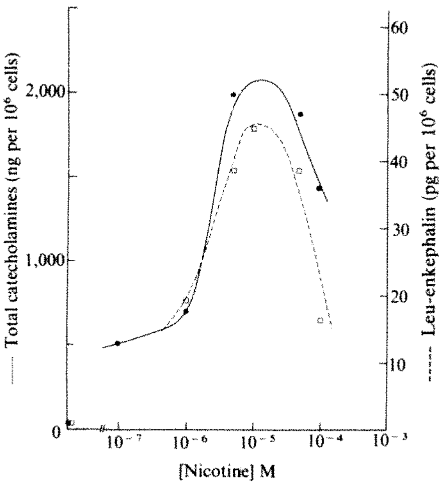


Fig. 2 Dose–response curves for release of endogenous catecholamines and Leu-enkephalin by nicotine. Adrenal paraneurones cultured on 14 Aclar hats were equilibrated for 15 min in KRH medium (containing 2.2 mM Ca^{2+} and 5 × 10^{-5} M Bacitracin), and this medium was then replaced by the same medium but containing the concentrations of nicotine indicated. At each concentration of agonist, two hats containing 5 × 10^5 cells per hat were pooled for analysis. After 5 min at 37 °C, the media were removed, acidified and chilled on ice and the cells collected for analysis of endogenous catecholamines and Leu-enkephalin as described in Table 1 legend.

This study was supported in part by a grant to B.G.L. from the Canadian MRC. J.R. was supported by the Institute National de la Santé et de la Recherche Medicale (France) and D.M.D. by a pre-doctoral fellowship from the Muscular Dystrophy Association of Canada. We thank Mr Joe Donahue for photography, Mrs Gwen Landrigan for typing the manuscript, and Dr R. P. Elde for providing the antisera used for the immunofluorescence studies.

Note added in proof: We have recently shown that Leu- and Met-enkephalin are co-stored within adrenaline (but not within noradrenaline) cells in the bovine adrenal medulla (B.G.L., R. Day, R. P. Elde and P. R. C. Howe, unpublished). A 'preliminary note' by S. M. Stine, H.-Y. T. Yang and E. Costa (*Neuropharmacology* **19**, 683–685 [1980]) indicates that Met-enkephalin-like immunoreactive material is released from isolated bovine chromaffin cells both spontaneously and in response to high concentrations of K^+ (56 mM) and acetylcholine (10^{-4} M).

Received 30 June; accepted 9 October 1980.

1. Hughes, J. *et al. Nature* **258**, 577–579 (1975).
2. Simantov, R. C. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2515–2519 (1976).
3. Bloom, F., Battenberg, E., Rossier, J., Ling, N. & Guilleman, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1591–1595 (1978).
4. Frederickson, R. C. A. *Life Sci.* **21**, 23–42 (1977).
5. Bayon, A., Rossier, J., Mauss, A., Bloom, F. E. & Iverson, L. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3503–3506 (1978).
6. Henderson, G., Hughes, J. & Kosterlitz, H. W. *Nature* **271**, 677–679 (1978).
7. Alumets, J., Hakansson, R., Sundler, F. & Chang, K. J. *Histochemistry* **56**, 187–196 (1978).
8. Hughes, J., Kosterlitz, H. W. & Smith, T. W. *Br. J. Pharmac.* **61**, 639–647 (1977).
9. Schultzberg, M. *et al. Acta physiol. scand.* **103**, 475–477 (1978).

10. Schultzberg, M. *et al. Neuroscience* **4,** 249–270 (1979).
11. Di Giulio, A. M. *et al. Neuropharmacology* **17,** 989–992 (1978).
12. Di Giulio, A. M., Yang, H. Y., Fratta, W. & Costa, E. *Nature* **278,** 646–647 (1979).
13. Schultzberg, M. *et al. Neuroscience* **3,** 1169–1186 (1978).
14. Viveros, O. H., Diliberto, E. J., Hazum, E. & Chang, K. J. *Molec. Pharmac.* **16,** 1101–1108 (1979).
15. Viveros, O. H., Diliberto, E. J., Hazum, E. & Chang, K. J. *Psychopharmacologia* **22,** 191–204 (1980).
16. Stern, A. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 6680–6683 (1979).
17. Clement-Jones, V., Lowry, P. J., Rees, L. H. & Besser, G. M. *Nature* **283,** 295–297 (1980).
18. Livett, B. G., Kozousek, V., Mizobe, F. & Dean, D. M. *Nature* **278,** 256–257 (1979).
19. Mizobe, F., Kozousek, V., Dean, D. M. & Livett, B. G. *Brain Res.* **178,** 555–566 (1979).
20. Lemaire, S., Lemaire, T., Dean, D. M. & Livett, B. G. *Nature* **288,** 303–304 (1980).
21. Livett, B. G. & Dean, D. M. *Neuropeptides* **1,** 3–13 (1980).
22. Rossier, J., Dean, D. M., Livett, B. G. & Udenfriend, S. *Life Sci.* (in the press).
23. Rossier, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **74,** 5162–5165 (1977).
24. Hökfelt, T., Johansson, D., Ljungdahl, A., Lundberg, J. M. & Schultzberg, M. *Nature* **284,** 515–521 (1980).
25. Miura, Y., Campese, V., De Quatro, V. & Meijer, D. *J. Lab. clin. Med.* **89,** 421–427 (1977).
26. Renzini, V., Brunori, C. A. & Valori, C. *Clin. chim. Acta* **30,** 587–594 (1970).
27. Fenwick, E. M., Fajdiga, P. B., Howe, N. B. S. & Livett, B. G. *J. Cell Biol.* **76,** 12–30 (1978).
28. Trifaro, J. M., Poisner, A. M. & Douglas, W. W. *Biochem. Pharmac.* **16,** 2095–2100 (1967).
29. Dean, D. M., Watson, L., Bray, G. M. & Livett, B. G. *Neurotoxins* (eds Chubb, I. W. & Geffen, L. B.) 261 (Adelaide University Union Press, 1979).
30. Wood, P. M. *Brain Res.* **115,** 361–375 (1976).
31. Bornstein, M. B. *Lab. Invest.* **7,** 134–137 (1958).
32. Furness, J. B., Costa, M. & Blessing, W. W. *Histochem. J.* **9,** 745–750 (1977).

Photolabelling of cholera toxin subunits during membrane penetration

Bernadine J. Wisnieski & John S. Bramhall*

Department of Microbiology and The Molecular Biology Institute, University of California, Los Angeles, California 90024

There has been much speculation about the mechanism by which cholera toxin exerts its effect on the cytoplasmic side of the membranes with which it interacts[1–4]. After the pentamer of B subunits (5B) binds to membrane receptors, particularly the monosialylganglioside GM_1, the disulphide-linked dimer A_1SSA_2 (which together with 5B constitutes the complete toxin) is thought to penetrate the membrane, perhaps through a channel formed by 5B and become reduced so that A_1SH units reach the cytoplasm and stimulate adenylate cyclase. Evidence for this mechanism is circumstantial[1,4,5]. If it is correct, a compound which will specifically label intramembranous sections of the toxin should label the channel-forming B subunits but not the channel-contained A_1 subunit. We have tested this prediction with a photoreactive glycolipid compound[6,7] and have obtained the opposite result. Therefore, we propose that only the A_1 subunit enters the membrane and we provide here data on the kinetics of that process.

The experimental protocol described is a direct and nonperturbing method for studying the dynamics of membrane penetration by proteins and other molecules. For simplicity, we used the Newcastle disease virus (NDV) system, earlier characterized with the photoreactive glycolipid probe[6], as a membrane target for cholera toxin because the viral envelope contains only three proteins, designated HN, F and M (ref. 8); surface neuraminidase activity resident in the HN protein makes NDV especially rich in GM_1 (ref. 8). The glycolipid probe 12-(4-azido-2-nitrophenoxy)-stearoyl[1-^{14}C]glucosamine (12APS-GlcN) is unique in that it will spontaneously insert into any membrane, and studies with a spin-labelled (nitroxide) counterpart suggest that the glucosamine moiety restricts it to the surface monolayer for an appreciable time after insertion[9]. The photoreactive group of 12APS-GlcN is located on the stearic acid chain of the amphiphilic compound. It reacts with vicinal compounds only when samples are exposed to UV light. Photolabelling is relatively nonspecific[7]. The probe labels coliphage $M13(f_1)$ coat protein only after incorporation into the bilayer of artificial vesicles[7]. Further, and most significantly, it labels only those residues of the protein which reside within the outer monolayer (ref. 10 and unpublished data). Chowdry and Westheimer[11] have reviewed membrane photolabelling studies. We have previously shown that cholera toxin can be photolabelled after binding to NDV charged with probe[12,13]. Here we analyse the kinetics of specific subunit insertion into the membrane.

Figure 1 shows the separation of proteins derived from cholera toxin, NDV and NDV with bound cholera toxin. The difference between the amount of toxin bound to NDV and the total toxin present (60 μg per sample) can be estimated by comparing lanes 1 and 3 and lanes 4 and 5. Typically, 3–5 μg of the toxin bound to NDV in the conditions used. Next to each stained gel the corresponding fluorogram shows proteins radiolabelled by the photoreactive probe (Fig. 1). Irradiation for 15 s allowed as much labelling of membrane proteins as irradiation for 6 min (data not shown); thus we conclude that most of the probe was photoactivated. Gel bands designated VP represent viral proteins whose photolabelling has been described elsewhere[6]. About 15–25% of the protein in band

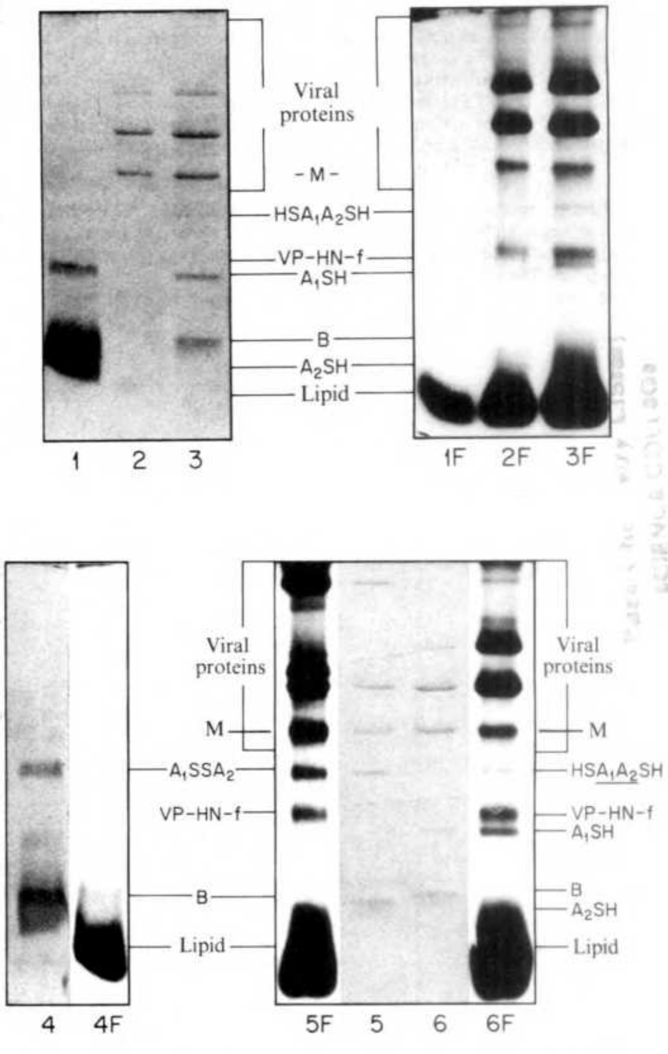


Fig. 1 Radiolabelling of NDV and cholera toxin proteins with the photoreactive probe 12APS-GlcN. Stained gel lanes are 1–6; corresponding fluorogram lanes are 1F–6F. Unreduced samples are in lanes 4 and 5. Lanes 1 and 4 contain cholera toxin (60 μg) labelled in solution and acid precipitated; lane 2 contains NDV (30 μg protein) labelled in suspension and pelleted. Lanes 3, 5, and 6 contain cholera toxin labelled after attachment to NDV for 1 min at 37 °C. Each sample was irradiated at 360 nm for 15 s. The amount of probe (50 mCi mmol^{-1}) per sample was 4×10^{-2} μCi. Total equilibration time with probe was 16 min at 37 °C for toxin alone and for NDV alone, and 15 min for NDV to which toxin was added 1 min before irradiation. Procedures as described in Table 1 legend.

---

* Present address: Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305.

Table 1 Photoreactive radiolabelling of cholera toxin subunits and viral proteins as a function of time and temperature of incubation

	Net c.p.m.						
Unreduced protein bands	0 °C, 15 min	37 °C, 30 s	37 °C, 1 min	37 °C, 3 min	37 °C, 15 min	NDV alone	Toxin alone
VP-M	209	174	196	210	192	205	0
A_1SSA_2	6	91	136	86	71	0	21
VP-HN-fragment	100	100	105	99	100	102	0
Total c.p.m.	315	365	437	395	363	307	21

Reduced protein bands	0 °C, 15 min	37 °C, 30 s	37 °C, 1 min	37 °C, 2 min	37 °C, 3 min	37 °C, 10 min	37 °C, 15 min	NDV alone	Toxin alone
$HS\underline{A_1A_2}SH$	2	23	35	27	23	20	19	0	4
VP-M	200	194	191	223	208	195	190	201	0
VP-HN-fragment	109	108	101	107	110	103	99	102	0
A_1SH	5	69	107	84	71	56	45	0	16
Total c.p.m.	316	394	434	441	412	374	353	303	20
$A_1SH + HS\underline{A_1A_2}SH$	7	91	142	111	94	96	64	0	20

Experiments were carried out under red safety lights until completion of electrophoresis to avoid activation of probe in unirradiated controls. Cholera toxin, purified[14] and stored at 6 mg ml^{-1} in 0.1 M sodium phosphate buffer, *p*H 7, was thawed and diluted to 1 mg ml^{-1} with 5 mM sodium phosphate-buffered saline containing 1 mM EDTA, *p*H 7.4 (PBSE). Two aliquots (0.6 ml) were spun at 149,000*g* for 30 min to remove any aggregates. The supernatant was used throughout. NDV strain HP-16, isolated as described elsewhere[6], was diluted to 450 μg protein per 9.4 ml PBS. The photoreactive probe 12APS-GlcN (50 mCi mmol^{-1}), synthesized as described elsewhere[15], was diluted to 90,000 c.p.m. per 10 μl of ethanol. In tube A, 150 μl of probe was dried and 50 μl of ethanol added and in tube B, 40 μl of probe was dried and 13 μl of ethanol added. The NDV suspension (9.4 ml) was added to tube A and 240 μl of toxin solution plus 2.5 ml PBSE added to tube B. Both tubes were incubated at 37 °C for 15 min. During this time, toxin solution was dispensed in 60-μl portions into 11 borosilicate tubes (6 × 50 mm), then 0.6 ml of NDV from tube A was added to each of these tubes, which were incubated for periods of time from 0.5 to 15 min before irradiation for 15 s at 366 nm with a mineral lamp (UV-Products, San Gabriel, California). Usually two samples were irradiated at each time point to provide reduced and non-reduced samples for gels. One unirradiated sample provided a control for background radioactivity. Two samples (0.6 ml) of NDV from tube A were incubated for an additional 15 min at 0 °C and at 37 °C before irradiation. Another two 0.6 ml samples of NDV were cooled to 0 °C before being added to two small tubes, each of which contained 60 μl of chilled toxin solution. One of these was irradiated after 1 min at 0 °C, the other after 15 min at 0 °C. The contents of tube B were divided into four 0.6-ml samples which were incubated at 0 °C for 1 min, 0 °C (15 min), 37 °C (1 min) and 37 °C for 15 min before irradiation. The entire experiment was repeated to provide a duplicate set of samples for analysis. Each time, 60-μl portions of toxin solution were added to 0.6 ml of PBSE and incubated for 30–60 min at 37 °C and centrifuged to determine how much toxin aggregated and could be pelleted without acid precipitation. This routine check was performed because it was found that commercial lyophilized toxin did not give reproducible results when used as a substitute for freshly isolated toxin (probably because it proved to be extremely difficult to completely solubilize). All samples were pelleted at 149,000*g* for 30 min on a Beckman SW50.1 rotor in buckets fitted with 0.7 ml adaptors. To each sample containing toxin with probe alone (from tube B), 50% trichloracetic acid (60 μl) was added after irradiation and the samples cooled to 0 °C before centrifugation. Pellets were boiled for 2 min in 40 μl of reducing or non-reducing SDS-buffer[16]. Electrophoresis and fluorography were carried out as described elsewhere[13]. Gel slices were oxidized to $^{14}CO_2$, which was collected and counted[6]. Top, data from non-reduced samples; bottom, reduced samples. Columns designated NDV alone and toxin alone refer to samples incubated independently with probe for 15 min at 37 °C, before irradiation. The sample of toxin alone (60 μg) was acid precipitated. The other columns represent components of toxin which pelleted with virus, estimated to be 5% of the 60 μg added. VP, viral protein; HN-f, HN-fragment. The $HS\underline{A_1A_2}SH$ peptide is the reduced but uncleaved $\underline{A_1A_2}$ peptide. Normally A_1SSA_2 completely reduces to A_1SH and A_2SH; however, in rapidly isolated cholera toxin preparations, not all the linear $\underline{A_1A_2}$ peptide is nicked between the two cysteines which form the disulphide bridge between A_1 and A_2.

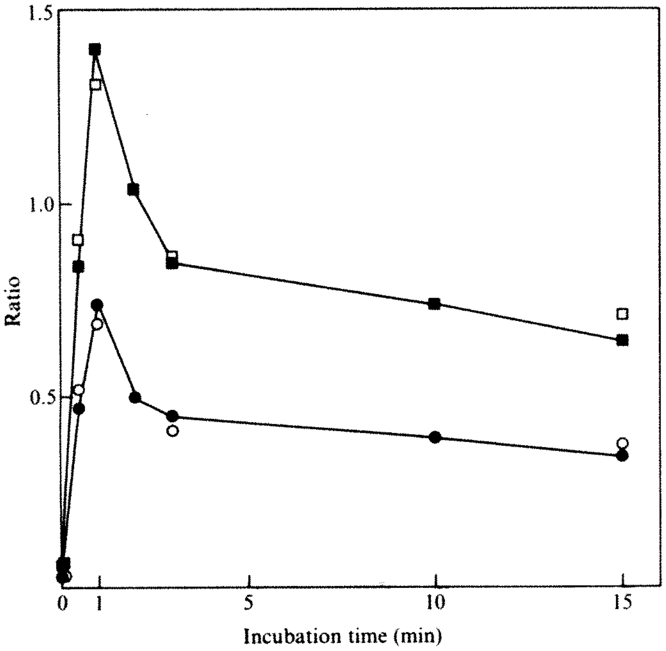


Fig. 2 The kinetics of photoreactive labelling of cholera toxin A subunits. Open symbols, ratios calculated from Table 1 (top); solid symbols, ratios calculated from Table 1 (bottom). Zero time points represent ratios calculated from 0°, 15′ samples in Table 1. Solid symbols, ratio of radiolabel in $A_1SH + HS\underline{A_1A_2}SH$ (unnicked) to radiolabel in viral protein M (●) and HN-f (■). Open symbols, ratio of radiolabel in A_1SSA_2 (nicked plus unnicked) to radiolabel in viral protein M (○) and HN-f (□). The ratio of radiolabel in HN-f to radiolabel in M was relatively constant (average 0.51; range 0.47–0.57).

A_1SSA_2 represents $\underline{A_1A_2}$ linear peptide which was not proteolytically cleaved during toxin purification[12] (the underlining of A_1A_2 is meant to distinguish this linear (uncleaved) peptide from the cleaved peptide A_1A_2). The rest of the band represents peptide which was cleaved between the disulphide bridge into discrete A_1 and A_2 subunits. The B subunit contains an internal disulphide bond and migrates with a larger apparent molecular weight when reduced. Reduced toxin samples show a faint band representing the reduced, but uncleaved, linear peptide $HS\underline{A_1A_2}SH$ (ref. 12), as well as A_1SH and B. The A_2SH subunit does not stain effectively. It migrates ahead of B and can be detected when ^{125}I-labelled toxin is electrophoresed[12]. In the fluorograms shown, the A_2SH peptide could not be resolved with any confidence from the diffuse band of radiolabelled lipids.

Other samples, including a series irradiated at defined times after the addition of toxin to NDV, were similarly electrophoresed. Bands representing viral proteins VP-M and VP-HN-fragment and toxin subunits (except A_2SH) were cut from stained gels after fluorography (because VP-HN-fragment does not stain (Fig. 1)), and oxidized to $^{14}CO_2$ which was trapped and counted. The level of radioactivity in the B subunit was not significantly above background in conditions tested (Fig. 1 and ref. 13). As previously demonstrated[6,13], unirradiated samples did not contain any radiolabelled proteins. Radioactivities of the pertinent bands are shown in Table 1. The labelling of A_1 plus A_2, that is, the net radioactivity associated with unreduced A_1SSA_2 (cleaved plus uncleaved), is expressed as a function of time and temperature of incubation with NDV before the 15 s irradiation. To standardize the data between gel lanes, raw data were later expressed as ratios within a particular lane. Radiolabelling of the two viral proteins chosen for ratio calculations did not vary with time or temperature or the presence of reducing agent, and therefore accurately reflect the amount of sample in each gel lane.

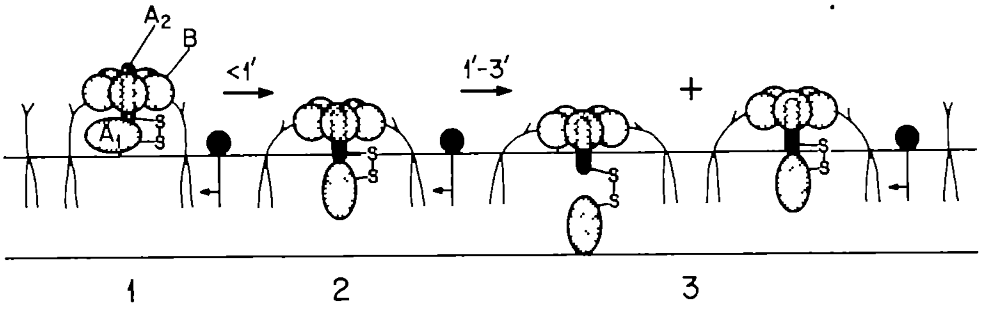


Fig. 3 Transmembrane kinetics of the active A_1 subunit of cholera toxin. Step 1, B subunits bind to ganglioside receptors (forked structures) Within 1 min of binding at 37 °C, A_1 insertion into the outer monolayer of the membrane is complete (Step 2). Within 2 min of insertion, A_1 equilibrates between the outer and inner monolayers (Step 3) The equilibrium achieved is undoubtedly a dynamic one, however, relative to the kinetics of the photolabelling reaction, the position of each A_1 is fixed. The small arrow represents the photoreactive group of 12APS–GlcN.

Data in Table 1 indicate that the A subunit of toxin can enter the vicinity of photogenerated reactive groups within seconds of binding to the viral membrane. To assess whether the radioactivity in the A_1SSA_2 band was due to labelling of A_1, A_2, or both, the ratios of counts in this band to the counts in both VP-M and VP-HN-f were plotted as a function of incubation time at 37 °C with NDV (Fig. 2, open symbols). The counts in A_1SH and in uncleaved $\overline{HSA_1A_2SH}$ were summed and expressed as a ratio to counts in $\overline{VP\text{-}M}$, and to counts in VP-HN-f (Fig. 2, solid symbols). Surprisingly, Fig. 2 shows that the ratios obtained as a function of time using total A_1SSA_2 (unreduced) counts and the ratios obtained using a formula in which 75–85% of A_2 is not included (due to reduction and dissociation of the cleaved fraction of A_1SSA_2 molecules) are not significantly different. From this we conclude that the majority of the radiolabel associated with the A_1SSA_2 band (unreduced) and the $\underline{SHA_1A_2SH}$ (reduced, uncleaved) is actually attached to the A_1 subunit of the toxin and not the A_2 portion.

The most striking feature of the data presented in Table 1 and in Fig. 2 is that the A component of toxin, more specifically the A_1 portion (from the rationale above), becomes increasingly accessible to photolabel within the first minute of incubation at 37 °C with NDV. During the next 2 min of incubation, radiolabelling of A_1 decreases by roughly 40% (34–41%). Between 3 and 15 min of incubation, very little change occurs. The accessibility of A_1 to probe after 15 min is 50% (46–54%) of that observed at 1 min. The rapid penetration and labelling of A_1 resemble that of complement component C9 on addition to vesicles containing C5b-8 (refs 17, 18 and unpublished data). The ratio of label in uncleaved toxin (HSA_1A_2SH) to label in A_1SH (calculated from data in Table 1) undergoes a slight increase with time; the significance of this remains to be determined. The kinetics of labelling of the A_1 subunit was verified qualitatively by exposing X-ray film to gels for various time periods before development. The first A_1-containing bands to produce silver grains in the film emulsion were always from the very early incubation times. Little or no labelling of any toxin subunit occurred when toxin and NDV were incubated together at 0 °C before irradiation, regardless of incubation time, even though NDV bound the same amount of toxin at this temperature as at 37 °C and displayed no change in its own labelling profile.

Values for the extent of labelling of the viral proteins and of the bound cholera toxin A_1 subunit (at the 1 min time point) were all at the level of 4 mol%. A similar value was obtained with M13 coat protein in vesicles[7]. This strongly implies that within 1 min of toxin addition and binding, most of the A_1 associated with virus entered the membrane envelope. If we assume that only 10% or 1% of the virally bound A_1 entered the membrane, the calculated level of A_1-specific labelling needs to be increased to 40 and 400 mol%, respectively, an unlikely occurrence unless an equivalent level of preference of probe for A_1 is proposed. Thus, we propose that most of the A_1 associating with virus actually penetrates the bilayer. We believe this to be the first such demonstration.

Our data are incompatible with the Gill[1] model for toxin entry because if the B pentamer formed a transmembrane channel, we would expect to label B but not A (which would be protected from the probe by residues of the B channel). The reverse was actually observed, and the detected entry of the enzymatically active A_1 subunit was found to be temperature-dependent, occurring at 37 °C (a temperature compatible with physiological activity), but not at 0 °C (a temperature compatible with binding but not activity).

Explanation of the kinetic data shown in Fig. 2 must answer two questions: (1) where does 50% of the A_1 entering the membrane go and (2) why does only 50% of the entering A_1 leave? One proposal (see Fig. 3), is based on the assumption that the probe remains restricted to the outer monolayer. This is a reasonable assumption because, as previously mentioned, all available data are consistent with an asymmetric distribution of such a compound (refs 9, 10 and unpublished data). Within 1 min of toxin binding, the A_1 subunit penetrates the outer monolayer of the membrane (step 1). The kinetics of A_1 entry between 0 and 1 min would reflect the kinetics of toxin binding to receptor[19] and subunit penetration. Between 1 and 3 min, the rapid decrease in A_1 labelling suggests that about half of the A_1 molecules (or one half of each molecule) leave the vicinity of probe within 2 min of entry (step 2). We visualize this as an equilibration of A_1 between the outer and inner monolayers until roughly one half is in each monolayer. The calculated half-time for the A_1 equilibration process (decrease in labelling) is of the order of 10–20 s. This time factor is not unreasonable for a diffusion-limited process in a viscous medium. At any time after 3 min, only 50% of the A_1 that entered the membrane remains accessible to probe, indicating very little change in A_1 distribution after 3 min (step 3).

By use of photoreactive probes to investigate the dynamics of toxin translocation in other systems, we can start to examine the specific roles of membrane composition, fluidity and asymmetry on protein traversal through membranes.

B.J.W. is the recipient of USPHS Research Career Development Award GM-00228. J.S.B. was a Fulbright-Hayes Postdoctoral Scholar. Research was supported by USPHS Grant GM-22240, the University of California Cancer Research Coordinating Committee and the Academic Senate.

Received 7 July, accepted 10 November 1980.

1 Gill, D. M. *Biochemistry* **15**, 1242–1248 (1976).
2. Sahyoun, N & Cuatrecasas, P. *Proc. natn Acad Sci. U S A* **72**, 3438–3442 (1975).
3 van Heyningen, S *Biochem J.* **168**, 457–463 (1977)
4 Tosteson, M T. & Tosteson, D. C. *Nature* **275**, 142–144 (1978)
5. Moss, J *et al Proc. natn Acad Sci U S A* **73**, 3480–3483 (1976)
6. Bramhall, J S, Shiflett, M A & Wisnieski, B J *Biochem. J.* **177**, 765–768 (1979)
7 Hu, V W. & Wisnieski, B. J. *Proc. natn Acad Sci U S A* **76**, 5460–5464 (1979)
8 Scheid, R & Choppin, P W *Virology* **57**, 475–490 (1974)
9. Wisnieski, B. J. & Iwata, K K *Biochemistry* **16**, 1321–1326 (1977).
10. Simon, P. & Wisnieski, B. J. *Fedn Proc* **39**, 2190 (1980)
11. Chowdry, V. & Westheimer, F. H. *A. Rev Biochem* **48**, 293–320 (1979)
12 Wisnieski, B J, Shiflett, M. A, Mekalanos, J & Bramhall, J S. *J. supramolec Struct* **10**, 191–197 (1979).
13. Wisnieski, B. J & Bramhall, J S *Biochem biophys Res Commun.* **87**, 308–313 (1979)
14 Mekalanos, J J, Collier, R. J & Romig, W R *Infect. Immun* **20**, 552–558 (1978)
15. Bramhall, J, Ishida, B. & Wisnieski, B. *J. supramolec. Struct* **9**, 399–406 (1978)
16 O'Farrell, P H *J. biol. Chem* **250**, 4007–4021 (1975)
17. Hu, V. W, Esser, A. F., Podack, E. R. & Wisnieski, B. J *Fedn Proc* **39**, 1971 (1980)
18. Hu, V W., Esser, A. F., Podack, E. R. & Wisnieski, B J *Proc. 4th int. Congr Immun Abstr.* No 15 2 07 (1980)
19 Cuatrecasas, P. (ed) *The Specificity and Action of Animal, Bacterial and Plant Toxins* (Chapman & Hall, London, 1977).

The magnitude of signal amplification by ligand-induced receptor clustering

Charles DeLisi

Laboratory of Theoretical Biology, National Cancer Institute,
National Institutes of Health, Bethesda, Maryland 20205

Lateral mobility of plasma-membrane proteins is now documented for a variety of cell systems. Evidence is accumulating that the molecular clustering allowed by such mobility may be important in the transduction of information at the cell surface[1–7]. Moreover, in a wide variety of systems, only a very small fraction of receptors need to be bound to induce a response[8,9], such as in the mast cell/basophil system where receptor clusters as small as dimers can induce a response[10] and only three or four such dimers are necessary for local degranulation[11]. As receptors are mobile, some clustering is expected even in the absence of ligand. As an approximate indication of the numbers involved, a consideration of constraints on translational entropy by spontaneous pair formation predicts[12] about 100 pairs on average for a cell with 500,000 receptors (typical of mast cells[13]), if a pair is defined as two receptors whose peripheries are within 5 Å. This estimate need not present a problem as it can be reduced by several orders of magnitude by postulating orientational requirements, repulsive potentials and/or more stringent constraints on tanslational entropy. Moreover, some cell preparations do show a slow Ca^{2+}-dependent background release[2] and this may, in fact, reflect spontaneous clustering and suggest that the control of activity is not stringent for those cells. An alternative view of the regulation of activity, however, which may have broad implications, is that receptors must be in proximity for some minimum amount of time for the transmission of biological information. This residency requirement may, for example, reflect the time required for the initiation of Ca^{2+} flow[14] or for receptors to change conformation[15] or for the pair to bind an integral membrane protein[16]. Whatever the origin, the quantitative predictions of the idea, which I develop here, suggest, for biologically relevant parameter values, an exceedingly sensitive control mechanism that can amplify the probability of an event by a factor of 10^{20} or more.

An estimate of the probability that a pair, having formed, will be active t s later is easily made if the pair separation times are exponentially distributed. Thus, if t_- is the mean time for a pair to separate by diffusion to some distance Δ beyond which it cannot transmit a signal, the fraction of pairs remaining within Δ for t s after pair formation is $\exp(-t/t_-)$. The fraction that remains within Δ long enough to transmit a signal is approximately

$$\nu_0 = \exp(-t_D/t_-) \tag{1}$$

where t_D is the minimum time the perimeters must be within Δ for transduction of a signal.

When receptors are cross-linked by ligand, pair separation will depend not just on diffusion but on the time required to break the ligand–receptor bond. If k_1 and k_{-1} are the reactive forward and reverse rate constants for ligand–receptor interaction respectively, then the effective rate constant for pair dissociation is[17,18]

$$k_r = \frac{k_- k_{-1}}{k_1 + k_-} \tag{2}$$

and the effective dissociation time, t_r, is

$$t_r = 1/k_r \tag{3}$$

In equation (2), the diffusive reverse rate constant, k_-, is related to the diffusive dissociation time by

$$k_- = 1/t_- \tag{4}$$

For diffusion-limited reactions ($k_1 \gg k_-$), equation (2) becomes

$$k_r = k_-/K_0$$

where

$$K_0 \equiv k_1/k_{-1} \tag{5}$$

is the intrinisic equilibrium constant of a ligand site for a receptor site. The ligand-induced amplification (A) of the transduction probability is defined as the ratio of the probability of transduction in the presence of ligand to that in its absence. Thus

$$A = \exp\left[-\frac{t_D}{t_-}\left(\frac{1}{K_0} - 1\right)\right] \tag{6}$$

Note that equation (6) is the ratio of two probabilities, each probability being obtained by division by the actual number of pairs. Hence the actual number of pairs does not enter equation (6). Had a different definition of amplification been adopted in terms of an increase in numbers of pairs (and not just their duration), the amplification would be somewhat larger than estimated below, by perhaps two orders of magnitude. It will become evident that the distinction between the two definitions is neither qualitatively nor quantitatively important. Equation (6) was chosen because of its relative simplicity.

K_0 will be large even if binding is weak. For example, a binding free-energy change of 5.5 kcal corresponds to an intrinisic equilibrium constant of 10^4 at room temperature. A plot of ammplification as a function of the ratio t_D/t_- for $K_0 = 10^4$ rises exponentially, reaching 10^{10} at a ratio of 25 and 10^{32} at a ratio of 75 (Fig. 1). Thus the transduction probability can be changed from essentially zero in the absence of ligand (for example, $\simeq 10^{-10}$ at $t_D/t_- = 25$) to close to one in its presence. The large amplifications are consequences of very low levels of spontaneous signal transmission in the absence of ligand.

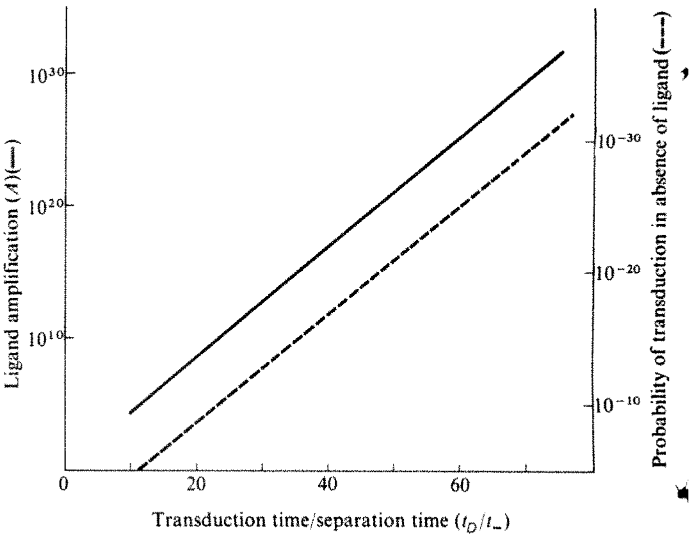


Fig. 1 Probability of transduction in the absence of ligand (– – –) and the amplification achieved (———) with a low-affinity ligand having an intrinsic equilibrium constant for the receptor of $K_0 = 10^4$.

Table 1 Mean time for a receptor pair to separate to a distance Δ

Δ(Å)	95	100	105	110	120	130
t_-(s)	1.3×10^{-6}	$5{\cdot}3\times10^{-6}$	1.2×10^{-5}	1.7×10^{-5}	4.3×10^{-5}	7.4×10^{-5}

Diffusion coefficient $= 3\times10^{-10}\,\text{cm}^2\,\text{s}^{-1}$, distance of closest approach = 90 Å.

How reasonable is the range of ratios for $t_D//t_-$ shown in Fig. 1? Transduction times are expected to be very rapid. For example, conformational changes in proteins in solution generally occur in $10^{-2}-10^{-3}$ s (ref. 19), though such changes might take somewhat longer in a membrane. Consequently, if transduction requires a conformational change, t_D would be of the order of milliseconds or longer. Similarly, experiments in which the flow of Ca^{2+} into a cell was followed indicate that uptake is 90% complete within 30 s, suggesting a time constant of seconds or less. Thus a transduction times of about 10^{-3} s would seem to be a reasonable lower bound. Amplification would then, according to Fig. 1, require diffusive dissociation times of $\leqslant 10^{-4}$ s.

The diffusive dissociation time can be estimated as follows. Consider circular receptor disks of diameter s with the plane of the disks coinciding with the plane of the membrane. The coordinate origin is fixed at the centre of one of the disks. Then, if the centre of a second disk is at $s \leqslant \rho \leqslant s+\Delta$, the mean time $W(\rho)$ required for its first arrival at $s+\Delta$ satisfies[20]

$$D\nabla^2 W(\rho) = -1 \tag{7}$$

where D is the sum of the diffusion coefficients of the disks. The solution to equation (7) subject to the boundary conditions

$$W(s+\Delta) = 0$$

$$\left(\frac{\mathrm{d}W}{\mathrm{d}\rho}\right)_{\rho=s} = 0$$

is

$$W(\rho) = \frac{(s+\Delta)^2 - s^2}{4D} + \frac{s^2}{2D}\ln\,(s/(s+\Delta)) \tag{8}$$

The mean time for a particle in the active region to escape is obtained by averaging equation (8) over all distances between s and $s+\Delta$. Thus,

$$t_- = \frac{2}{(s+\Delta)^2 - s^2}\int_s^{s+\Delta} \rho W(\rho)\,\mathrm{d}\rho \tag{9}$$

Diffusive dissociation times were calculated according to equation (9) as a function of Δ (Table 1), using a receptor diameter of 90 Å and a diffusion coefficient of $3\times10^{-10}\,\text{cm}^2\,\text{s}^{-1}$, numbers typical of the mast cell system[21,22]. The calculation is not of course intended to be precise, but is presented here only to provide perspective for evaluating the plausibility of the range of ratios in Fig. 1.

The above analysis assumes that when receptors are cross-linked, they are always within the active radius. However, this assumption is not compatible with molecular flexibility. Suppose, for example, that the doubly bound ligand were fully flexible and its maximum extension exceeded Δ. Then within the framework of the residency-time requirement, the ligand would not be sensed by the cell and the amplification would be one. The reason is that if the ligand is so flexible that it does not affect the diffusion of one receptor relative to its partner, the time to escape from the active region will not be changed; only the rate of entry into the region will change. A fully flexible molecule is, however, unrealistic.

For molecules in solution, some fluctuation in distances between groups is expected, even if the intervening chain is short. However, when both groups are constrained to lie on a cell surface as a result of bonding to receptors, the ability of the intervening backbone bonds to rotate is severely restricted—a fact easily verified by examination of molecular models. The chain is then effectively frozen in whatever configuration it was in when the second bond formed. From this point of view, the distances between reactive groups on a ligand, averaged over the collection of multiply bound ligands, will reflect the distribution of distances between groups when the ligand is free, even though the distance between groups on any particular molecule does not fluctuate. But only that fraction of multiply bound ligands having configurations constraining the peripheries of the receptors to be within Δ will be recognized. Thus the reduction in the concentration of active pairs, at a fixed ligand concentration, will be the main effect of flexibility and the amplification factor given by equation (6) remains essentially unchanged. Experiments using stiff chains or using flexible chains whose dimensions have been determined independently, for example by energy transfer, may elucidate the validity of these remarks.

Finally, note that the form of equation (6) is not necessarily limited to systems that rely on clustering. For example, if the step subsequent to binding were a conformational change, then t_- could be interpreted as the natural lifetime of the active conformation in the absence of ligand. In the presence of ligand, the lifetime would be k_r, the dissociation time of the ligand. The amplification would then be $\exp[k_r - k_-)t_D] \simeq \exp(k_r t_D)$, and the fraction of receptors remaining active for t_D, $\exp(-k_r t_D)$. As long as $k_r t_D$ is much less than one, changes in k_r will not affect the fraction of bound ligands that are active. As $k_r t_D$ becomes greater than one, however, activity should drop rapidly. Somewhat similar considerations[23] were used to explain the precipitous drop in cyclic AMP production as the affinity of hyponeurophyseal hormones for their receptors decreased, while the total bound concentration remained constant.

In conclusion, the molecular complex mediating information transfer, whether it be a receptor cluster or a particular receptor conformation, is assumed to be present at all times, but intrinsically unstable with a decay time that is short compared with the time required for information transfer. Signals in these circumstances are too low to be distinguished from background noise. Ligands, by stabilizing an active complex so that its decay time exceeds the time required for information transfer, selectively amplify particular signals. Numerical estimates indicate that such amplification can be very large.

I thank Drs Henry Metzger and Alan Perelson for reading and commenting on the manuscript.

Received 14 July; accepted 10 November 1980.

1. Siraganian, R. P., Hook, W. A. & Levin, B. B. *Immunochemistry* **12,** 149–157 (1975).
2. DeLisi, C. & Siraganian, R. P. *J. Immun.* **122,** 2286 (1979).
3. Dembo, M., Goldstein, B., Sobotka, A. K. & Lichtenstein, L. M. *J. Immun.* **122,** 518 (1979).
4. Fanger, M. W., Hart, D. A. Wells, J. V. & Nisonoff, A. *J. Immun.* **105,** 1484 (1970).
5. Archer, B. G. & Krakauer, H. *Biochemistry* **16** (1977).
6. Schechter, Y., Hernaez, L., Schlessinger, J. & Cuatrecasas, P. *Nature* **278,** 835 (1979).
7. Kahn, C. R., Baird, K. L., Jarrett, D. B. & Flier, J. S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4209 (1978).
8. Birnbaumer, L., Pohl, S. L. & Kaumann, A. J. *Adv. Cylic Nucleotide Res.* **4** (1975).
9. Schlessinger, J. in *Physical Chemistry of Cell Surface Events in Cellular Regular* (eds Delisi, C. & Blumenthal, R.) (Elsevier, Amsterdam, 1979).
10. Segal, D. M., Taurog, J. & Metzger, H. *Proc natn. Acad. Sci. U.S.A.* **74,** 2993 (1977).
11. DeLisi, C. & Siraganian, R. *J. Immun.* **122,** 2281 (1979).
12. DeLisi, C. *Q. Rev. Biophys.* (in the press).
13. Metzger, H. & Bach M. K. in *Immediate Hypersensitivity* (ed. Bach, M. K.) 561 (Dekker, New York, 1978).
14. Forman, J. C., Hallet, M. B. & Morgar, J. L. *J. Physiol., Lond.* **271,** 193 (1977).
15. Hiedman, T. & Changeux, J. P. *A. Rev. Biochem.* **47,** 317 (1978).
16. Bourguignon, L. Y. W. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5131 (1977).
17. Eigen, M. in *Quantum Statistical Mechanics in the Natural Sciences* (eds S. L. Minz, & Weidermayer, S. M.) (Plenum, New York, 1974).
18. Bell, G. I. *Science* **200,** 618 (1978).
19. Hammes, G. *Adv. Protein Chem.* **23,** 1 (1968).
20. Berg, H. C. & Purcell, E. M. *Biophys. J.* **20,** 193 (1977).
21. Metzger, H. *Immun. Rev.* **41,** 186 (1978).
22. Schlessinger, J. Webb, W., Elson, E. L. & Metzger, H. *Nature* **264,** 550 (1976).
23. Bergman, R. & Hechter, O. *J. biol. Chem.* **253,** 3238 (1978).

Evidence for a divided genome in bean golden mosaic virus, a geminivirus

S. Haber, M. Ikegami, N. B. Bajet & Robert M. Goodman

Department of Plant Pathology, University of Illinois, Urbana, Illinois 61801

Geminiviruses[1,2], a recently discovered group of plant viruses, are the only viruses containing circular single-stranded DNA (ssDNA) found in eukaryotic organisms. Their ssDNA, which is less than half the size (molecular weight (M_r) 7–8 × 10^5) (refs 3–6) of that of other ssDNA viruses[7,8], is the smallest yet found in autonomously replicating viruses. Geminivirus DNA is encapsidated in paired isometric particles whose monomer units measure ~18 nm. This unique capsid morphology was accounted for by an early hypothesis of a functionally divided genome. Among the RNA plant viruses, there are several groups that carry necessary genetic information on more than one nucleic acid species. In the case of geminiviruses, two nucleic acid components were postulated to be carried in each paired particle[9,10]. Analyses of the physical composition of geminiviruses showed, however, that there is only one DNA molecule in each paired particle[6,11,12]. We have now investigated the DNA of a geminivirus using restriction endonucleases and a biological assay of infectivity and report here evidence for a divided genome in bean golden mosaic virus (BGMV). Two different nucleotide sequences are found and both seem to be required for infectivity.

BGMV was maintained by sap transmission in *Phaseolus vulgaris* L. ('Top Crop') kept in simulated humid tropical conditions in a growth chamber[11,13] and purified as described elsewhere[13,14]. Viral DNA labelled with ^{32}P *in vivo* was obtained by collecting systemically infected trifoliolate leaves from plants 5 days after inoculation of primary leaves and feeding with ^{32}P-orthophosphate (1 mCi per leaf) through the cut ends of petioles. The leaves were then maintained for 4 days in a Petri dish by feeding a balanced inorganic salt solution through the petioles.

BGMV DNA is resolved by polyacrylamide gel electrophoresis in 7 M urea into two components which comprise a covalently closed circular molecule (M_r 8 × 10^5) and a slightly shorter linear molecule[5]; previous evidence suggests that the linear molecule is derived from (or possibly is a precursor of) the circular form[11]. We separated ^{32}P-labelled circular and linear molecules by gel electrophoresis and recovered the DNA by electro-elution[15] from excised gel slices. The DNA was then digested separately with restriction endonucleases *Hha*I and *Hae*III which have sequence-specific activity on the ssDNA of bacterial viruses[16,17]. DNA from ΦX174 digested with *Hae*III was used as a standard.

Digestion of linear or circular BGMV DNA with *Hha*I yielded fragments of the same size and relative electrophoretic mobility (data not shown), confirming our previous conclusion that circular and linear forms of BGMV DNA contain the same nucleotide sequences[11]. Calculation of the length of each BGMV DNA fragment present in complete *Hha*I digests of circular BGMV DNA revealed that the total length of these fragments (5,380 nucleotides) was about twice that expected (2,510 nucleotides) from electron-microscopic measurements of the DNA (ref. 5) (Table 1).

We also used a virus-specific double-stranded (ds) DNA, which we consider to be a replication intermediate, isolated from ^{32}P-labelled, BGMV-infected leaves. This contains the complementary strand to viral DNA and both strands are of

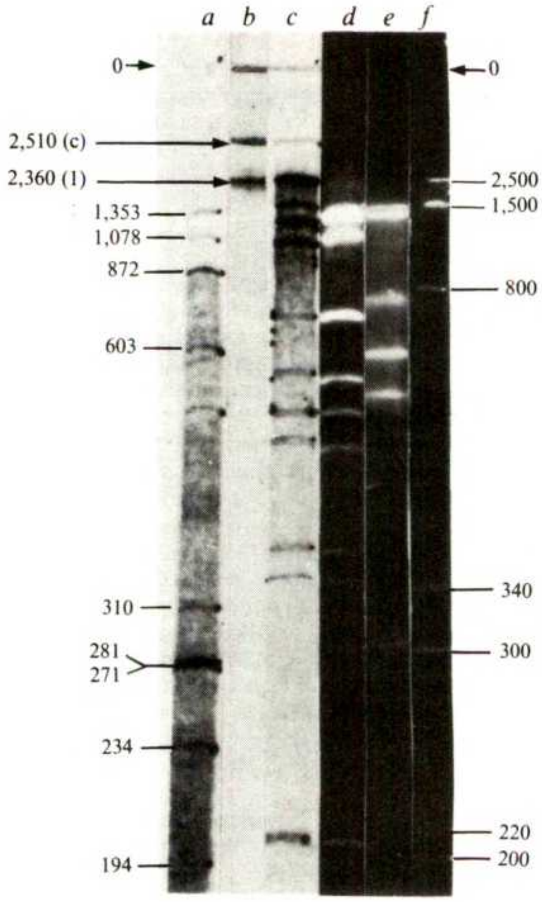


Fig. 1 Analysis of BGMV DNA by restriction endonucleases. *a*, 6 μg of ΦX174 viral ssDNA was incubated with 70 units of *Hae*III for 16 h at 37 °C in 50 mM Tris-HCl, *p*H 7.5 (120 μl), 5 mM $MgCl_2$ and 0.5 mM dithiothreitol. *b*, 4 μg of BGMV viral DNA was incubated as in *a*, but without restriction enzyme: c, circular DNA; l, linear DNA. *c*, 4 μg of BGMV viral DNA was incubated with 120 units of *Hha*I for 16 h at 37 °C in 50 mM Tris-HCl (*p*H 8.0, 500 μl) and 5 mM $MgCl_2$. *d*, 5,000 c.p.m. of ^{32}P-BGMV dsDNA, with 5 μg of ΦX174 RF DNA as carrier, was incubated with 20 units of *Hha*I as in *c*. *e*, 5,000 c.p.m. of ^{32}P-BGMV dsDNA with 5 μg of ΦX174 RF DNA as carrier was incubated with 20 units of *Hae*III as in *b*. *f*, 2,000 c.p.m. of ^{32}P-M13 RF DNA, prepared by the procedure of Messing (unpublished protocol), with 5 μg of ΦX174 RF DNA as carrier, was incubated as in *e*. After incubation, samples were phenol/chloroform extracted and the DNA ethanol-precipitated[15], then redissolved in 10 μl of sterile distilled water. Immediately before application to 4% polyacrylamide (acrylamide/bisacrylamide, 22:1) 7 M urea slab gels (0.75 mm × 13 cm × 14 cm)[15], each sample was added to 6 μl of a solution containing 50% sucrose, 30% urea, 0.2% xylene cyanol, 0.2% bromophenol blue and heated at 90 °C for 1 min. The electrophoresis buffer was 50 mM Tris-borate (*p*H 8.3), 1 mM EDTA. Electrophoresis was at 200 V until the xylene cyanol dye reached the bottom of the gel. Gels *a*, *b* and *c* were stained in the dark for 30 min with 0.005% Stains-All in 50% formamide, then destained in distilled water under light. For autoradiography, X-ray film (DuPont Cronex) was exposed to dried gels (*d*, *e* and *f*) in intensifying screen cassettes at −80 °C for ~4 d before developing. Arrows indicate the positions of restriction fragments from *Hae*III digests of ΦX174 viral DNA (left) and *Hha*I digests of M13 RF DNA (right); the numbers indicate approximate number of nucleotides (ΦX174) or nucleotide pairs (M13) of the fragments. Interpolated values for BGMV viral and dsDNA restriction fragments are given in Table 1. The top of the gel is indicated by the arrow labelled 0.

0028-0836/81/040324—03$01.00

Table 1 Restriction fragment sizes of ΦX174 ssDNA, BGMV ssDNA, BGMV dsDNA and M13 RF DNA shown in Fig. 1, in order of increasing electrophoretic mobility

DNA	Treatment	Fragment sizes (nucleotides or nucleotide pairs)	Remarks
ΦX174 ss	*Hae*III	1,353; 1,078; 872; 603; 310; 281; 271; 234; 194*	Sizes of these fragments were computed from the nucleotide sequence given in ref. 18
BGMV ss	Untreated	2,510 (circular); 2,360 (linear)	Sizes based on electron microscope measurements[5]
BGMV ss	*Hha*I	(2,510)†; (2,360); (1,650); 1,230; 1,050; (950); 660; 570; 520; 475; 365; 335; 175	Sizes based on interpolation using ΦX174 ssDNA *Hae*III fragments as standards. Sum of length of fragments in complete digests is 5,380 nucleotides
BGMV ds	*Hha*I	1,350; 1,030; 650; 520; 460; 420; 340; 320	Sizes based on M13 RF DNA *Hae*III fragments as standards. Sum of fragment lengths is 5,280 nucleotide pairs
BGMV ds	*Hae*III	1,480; 1,380; 750; 585; 510; 290	Sizes based on M13 RF DNA *Hae*III fragments as standards. Sum of fragment lengths is 4,995 nucleotide pairs
M13 ds	*Hae*III	2,500; 1,500; 800; 340; 300; 220; 200	Sizes of these fragments are from values given in a laboratory protocol (J. Messing, unpublished)

* Fragments 118 and 72 nucleotides long run off the gel in conditions used in our experiments (see Fig. 1 legend). Identity of these fragments was verified by comparison with *Hae*III restriction products of ΦX174 RF DNA. Other fragments in Fig. 1*a*, some of which occur as discrete bands, are due to the presence of ~20% by weight linear ssDNA in the ΦX174 ssDNA preparation.

† Values in parentheses are for DNA fragments produced by incomplete digestion of the substrate, as indicated by their absence in digests of BGMV dsDNA by the same enzyme, and by results of prolonged incubation of ssDNA with larger amounts of enzyme (not shown).

approximately unit length in relation to the circular viral DNA (M.I. and R.M.G., unpublished). The fragments produced by *Hha*I digestion of BGMV dsDNA corresponded exactly in number and size to those produced by complete *Hha*I digestion of circular viral ssDNA (Fig. 1).

Digestion of BGMV dsDNA with *Hae*III yielded six fragments with a total length of 4,995 nucleotide pairs (Table 1), confirming the results obtained with *Hha*I.

Attempts to analyse BGMV ssDNA by digestion with *Hae*III were not successful. *Hae*III digestion yields the same number and sizes of fragments from ΦX174 ssDNA as are found in *Hae*III digests of ΦX174 replicative form (RF) DNA[16]. The *Hae*III restriction fragments of circular BGMV viral DNA, however, were very small (20–100 nucleotides) and did not form discrete bands (data not shown). The nucleotide ratio of ^{32}P-labelled BGMV DNA is 32% A:33% T:22% G:13% C (unpublished results); thus it is unlikely that the *Hae*III recognition sequence (5′GG↓CC3′) is as numerous in such a small DNA (with G+C = 35%) as the restriction analysis suggested. Results with BGMV dsDNA confirm this. We cannot explain the apparently anomalous behaviour of *Hae*III with BGMV ssDNA.

These results indicated that BGMV may contain a divided genome with ssDNA components of the same physical size but different nucleotide sequence. The same result could be expected, however, if two related but distinct strains of the virus were present. We tested the divided genome hypothesis by investigating the dilution kinetics of infection by BGMV DNA in bean cell protoplasts. Protoplasts were prepared from leaf mesophyll cells of healthy beans (N.B.B and R.M.G., unpublished results), inoculated with different amounts of BGMV DNA, and infected cells scored after 24 h by immunofluorescence. The slope of the infectivity dilution plot (Fig. 2) was closer to two than one.

Part of the initial evidence that led to the discovery of divided genomes in other plant virus groups were infectivity dilution curves whose slopes deviated significantly from unity[19,20]. Our results now indicate that BGMV and possibly other geminiviruses may also contain a divided genome. This result extends to ssDNA viruses the interesting prevalence of divided genomes among the plant viruses and we believe it is the first example of a divided genome in a DNA virus. We conclude that the BGMV DNA genome is encapsidated in paired particles of two classes which, according to all present evidence, differ only in the sequence of the circular DNA molecule they contain. If this hypothesis is correct, the infectivity dilution kinetics of the virus particles should also approach a value of two; unfortunately, we have not found an experimentally useful local lesion host for BGMV and attempts to infect protoplasts by inoculation with whole virus have been unsuccessful.

This discovery is expected to complicate further studies on the molecular biology of BGMV. To sequence or map gene functions on BGMV DNA, we must be able to resolve two DNA

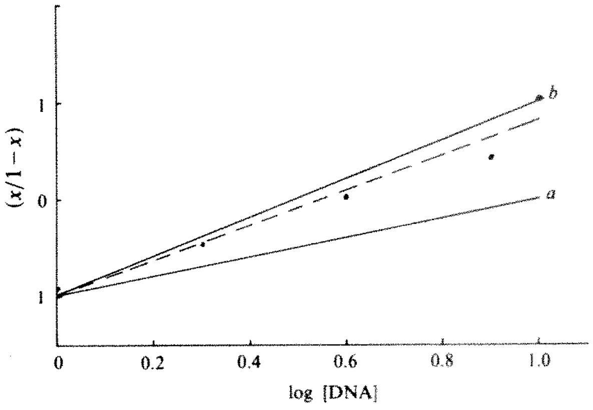


Fig. 2 Relationship between concentration of BGMV DNA in the inoculum and proportion of inoculated bean cell protoplasts infected. Plots shown are data from one (expt. 3) of five separate experiments (---) and theoretical plots (—) for *a*, one-hit and *b*, two-hit kinetics. Protoplasts were isolated from the mesophyll of Top Crop bean leaves and inoculated with various amounts of DNA (data are in μg ml^{-1}). Infected cells were scored by an indirect fluorescent-antibody method using anti-BGMV rabbit serum and fluorescein-labelled anti-rabbit γ-globulin goat serum. Data were corrected for the number of living cells determined by phenosaffranin staining[21] and then transformed to logit values[22] ($\log x/(1-x)$) where x is the proportion of living cells infected. The logit transformation was used to correct for the decreased availability of infectable cells at higher DNA concentrations. Data from five experiments were plotted and the lines drawn by linear regression analysis. The slopes, correlation coefficients (r) and the probability level (P) at which r values are significant are shown below:

Expt	1	2	3	4	5
Slope	1.99	1.71	1.71	1.78	1.78
r	0.97	0.96	0.98	0.99	0.97
P	0.01	0.05	0.01	0.01	0.01

species which are indistinguishable by size but differ in nucleotide sequence. Perhaps the best approach would be to find restriction endonucleases that will digest at one site in only one of the two sequences and then to clone that sequence out of a mixture of the two using recombinant DNA techniques.

We acknowledge the collaboration and advice of Dr K. S. Browning, and thank Drs C. L. Niblett and R. W. Blakesley for discussions and Dr P. D. Shaw for comments on an earlier draft. Research was supported by a grant (7800549) from the USDA Competitive Research Grants Office and from the Illinois Agricultural Experiment Station (project 68-366).

Received 15 August; accepted 13 November 1980.

1. Goodman, R. M. in *Handbook of Plant Viral Diseases* (ed. Kurstak, E.) (North Holland/Elsevier, Amsterdam, in the press).
2. Matthews, R. E. F. *Intervirology* **11,** 133–135 (1979).
3. Goodman, R. M. *Nature* **266,** 54–55 (1977).
4. Harrison, B. D. *et al. Nature* **270,** 760–762 (1977).
5. Reisman, D., Ricciardi, R. P. & Goodman, R. M. *Virology* **97,** 388–395 (1979).
6. Francki, R. I. B., Hatta, T., Boccardo, G. & Randles, J. W. *Virology* **101,** 233–241 (1980).
7. Ward, D. C. & Tattersall, P. *Replication of Mammalian Parvoviruses* (Cold Spring Harbor Laboratory, New York, 1978).
8. Denhardt, D. T., Dressler, D. & Ray, D. S. *The Single-Stranded DNA Phages* (Cold Spring Harbor Laboratory, New York, 1978).
9. Bock, K. R., Guthrie, E. J. & Woods, R. D. *Ann. appl. Biol.* **77,** 289–296 (1974).
10. Galvez, G. E. & Castano, M. *Turrialba* **26,** 205–207 (1976).
11. Goodman, R. M., Shock, T. L., Haber, S., Browning, K. S. & Bowers, G. R. Jr *Virology* **106,** 168–172 (1980).
12. Hatta, T. & Francki, R. I. B. *Virology* **92,** 428–435 (1979).
13. Goodman, R. M., Bird, J. & Thongmeearkom, P. *Phytopathology* **67,** 37–42 (1977).
14. Goodman, R. M. & Bird, J. *AAB/CMI Descriptions of Plant Viruses* No. 192 (Commonwealth Mycological Institute, Kew, 1978).
15. Browning, K. S. thesis, Univ. Illinois (1980).
16. Blakesley, R. W. & Wells, R. D. *Nature* **257,** 421–422 (1975).
17. Horiuchi, K. & Zinder, N. D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2555–2558 (1975).
18. Sanger, F. *et al. J. molec. Biol.* **125,** 224–246 (1978).
19. Fulton, R. W. *Virology* **18,** 466–485 (1962).
20. Van Vloten-Doting, L., Kruseman, J. & Jaspars, E. M. *Virology* **34,** 728–737 (1968).
21. Widholm, J. *Stain Technol.* **47,** 189–194 (1972).
22. Van der Planck, J. E. *Plant Disease: Epidemics and Control* (Academic, New York, 1963).

Plasmid R1 DNA replication dependent on protein synthesis in cell-free extracts of *E. coli*

Ramon Diaz & Kurt Nordström

Department of Molecular Biology, Odense University,
DK-5230 Odense M, Denmark

Walter L. Staudenbauer

Max-Planck-Institut für Molekulare Genetik, Abt. Schuster,
D-1000 Berlin 33, FRG

Bacterial plasmids serve as model systems for studying the regulation of DNA replication[1]. Elucidation of the molecular mechanisms involved in plasmid DNA synthesis requires the development of efficient cell-free plasmid replication systems. Such *in vitro* systems have previously only been described for Col E1-type plasmids[2,3] and for the R6K plasmid[4,5]. Here we report that extracts of *Escherichia coli* can carry out the complete replication of miniplasmids derived from the antibiotic-resistance plasmid R1. This R1 replication system differs from the previously described ColE1 and R6K systems in its strict dependence on DNA-directed protein synthesis. We believe this to be the first report of the functional coupling of the three fundamental reactions of genetic information transfer (transcription, translation and replication) in a cell-free system.

Table 1 Requirements of the system

Omissions	dCMP incorporated (pmol)
None	94.9
$MgCl_2$	0.8
dTTP	5.1
ATP	2.1
CTP, GTP, UTP	3.5
Creatine phosphate	1.0
Cyclic AMP	36.7
Amino acids	40.3

Extract was prepared from C600 (pKN177) and supplemented with exogenous pKN177 DNA (30 μg ml^{-1}). The complete system is the standard reaction mixture described in Fig. 1 legend with an additional 0.5 mM of each of the 20 amino acids. [α-^{32}P]dCTP (1,000 c.p.m. pmol^{-1}) was used as radioactive label. Reaction mixtures (25 μl) were incubated in duplicate at 30 °C for 60 min.

To develop an *in vitro* replication system for the R1 plasmid, we used the mini-R1 plasmids pKN177 and pKN182, which had been spontaneously generated from the R1 Cop mutant pKN104 (ref. 6). Plasmid pKN182 has a molecular weight of 3.9×10^6 and carries only the replication region of the parent plasmid, whereas pKN177, having a molecular weight of $\sim 9 \times 10^6$, also contains the ampicillin-resistance gene of R1. The Cop phenotype is caused by a single base substitution in the replication region[7], which prevents the expression of a *trans*-acting

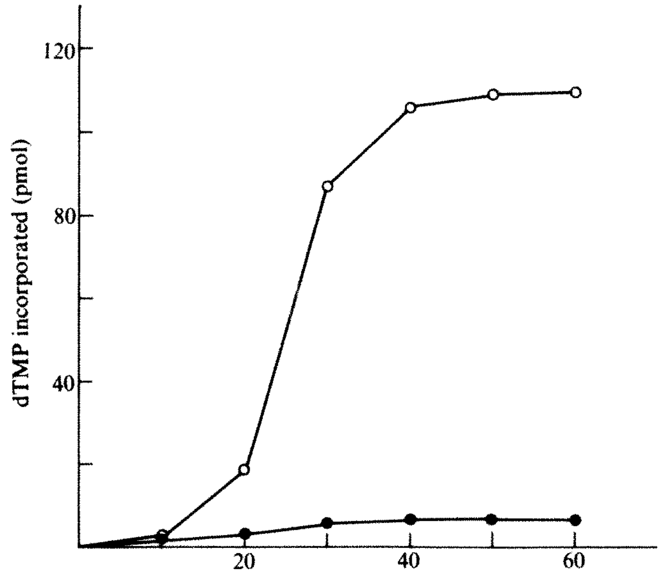


Fig. 1 Kinetics of pKN177 DNA synthesis. A cell-free extract was prepared from *E. coli* C600 (pKN177) by the freeze–thaw method[8] in the presence of 1 mM EDTA and 1 mM dithiothreitol. Standard reaction mixtures contained final concentrations of 50 mM HEPES *p*H 8.0, 100 mM KCl, 11 mM magnesium acetate, 15 mM creatine phosphate, 0.1 mg ml^{-1} creatine kinase, 2 mM ATP, 0.4 mM each of CTP, GTP and UTP, 0.05 mM NAD, 0.05 mM cyclic AMP and 0.025 mM each of dATP, dCTP, dGTP and ^{3}H-dTTP (500 c.p.m. pmol^{-1}). Reaction mixtures (500 μl) containing 250 μl extract (7.5 mg protein) were incubated at 30 °C without (●) or with (○) addition of exogenous pKN177 DNA (30 μg ml^{-1}). At the times indicated, 25-μl samples were removed and assayed for acid-insoluble radioactivity.

0028-0836/81/040326—03$01.00

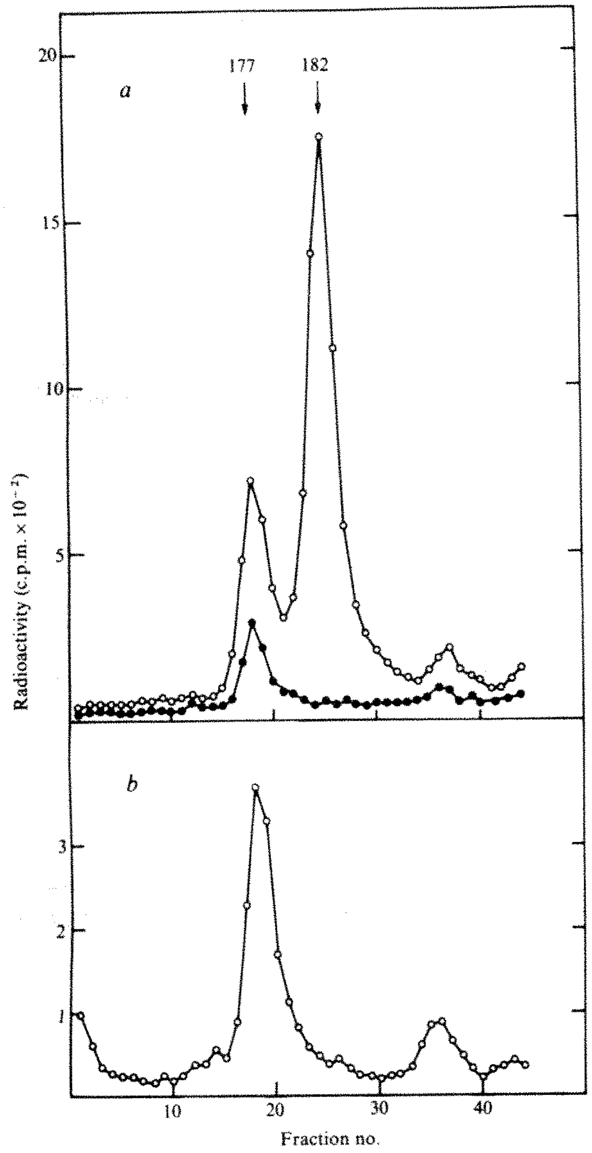


Fig. 2 Alkaline CsCl gradient centrifugation of *in vitro* synthesized DNA. Extracts were incubated for 60 min at 30 °C in standard reaction mixtures (150 µl) without or with addition of exogenous plasmid DNA (30 µg ml^{-1}). Reactions were stopped by addition of EDTA (50 mM) and the DNA was purified by gel filtration through a Sepharose 6B solumn. Aliquots were layered on preformed 4.2-ml alkaline CsCl gradients (density 1.2–1.4 g ml^{-1} in 0.2 M NaOH) and centrifuged for 35 min in a Beckman SW 50.1 rotor at 5 °C. Fractions (100 µl) were collected from the bottom of the tubes and assayed for radioactivity. Centrifugation is from right to left. The arrows indicate the positions of supercoiled pKN177 and pKN182 DNA, respectively. *a*, C600 (pKN177) extract incubated without (●) or with (○) addition of pKN182 DNA; *b*, C600 extract incubated with addition of pKN177 DNA.

negative control function[6]. Due to their reduced size and increased copy number (~25 copies per chromosome), these mini-R1 plasmids offer experimental advantages for setting up an *in vitro* system for R1 DNA replication.

Plasmid DNA synthesis was first studied in extracts from the plasmid-carrying *E. coli* strain C600 (pKN177) with the assumption that mini-R1 replication may require a plasmid-coded replication protein. We found that extracts prepared by freeze–thaw lysis of lysozyme-treated cells can catalyse a low but significant amount of dTMP incorporation into endogenous plasmid DNA. Furthermore, a marked stimulation of DNA synthesis was observed on addition of exogenous pKN177 DNA. The kinetics of dTMP incorporation with or without addition of plasmid DNA is shown in Fig. 1. It can be seen that there is a lag period of 10 min before the onset of DNA synthesis and that the incorporation levels off after 40 min of incubation. Addition of pKN177 DNA stimulates the total incorporation

Table 2 Replication capacity of extracts prepared from plasmid-carrying and plasmid-free strains

	Plasmid DNA added		
Source of extract	None	pKN177	ColE1
C600 (pKN177)	14.8	138.1	182.9
C600	0.0	9.5	163.5

Standard reaction mixtures (25 µl) included ^{3}H-dTTP (500 c.p.m. pmol^{-1}) and 0.5 mM of each of the 20 amino acids. Exogenous pKN177 DNA (30 µg ml^{-1}) or ColE1 DNA (50 µg ml^{-1}) were added as indicated. Incubations were performed in duplicate at 30 °C for 60 min. Values shown are dTMP incorporated (pmol).

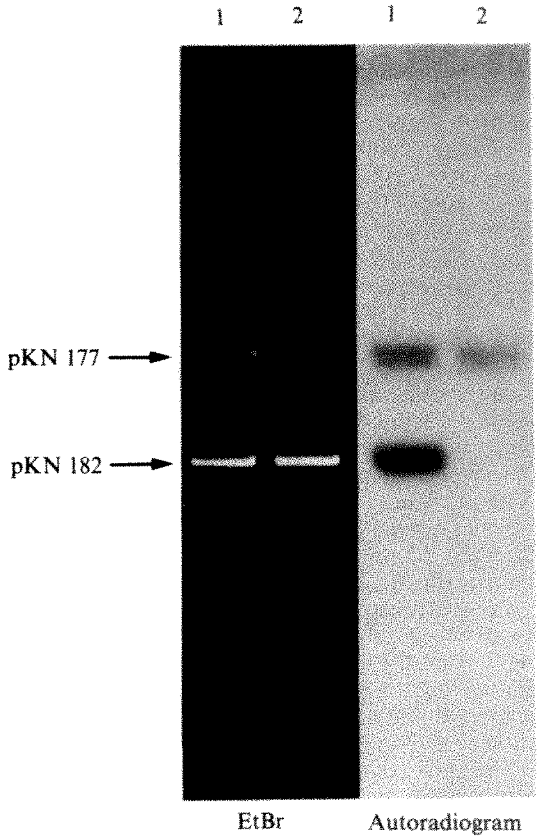


Fig. 3 Effect of rifampicin on plasmid DNA replication. Extract from C600 (pKN177) was incubated for 60 min at 30 °C in standard reaction mixtures (150 µl) containing [α-^{32}P]-dCTP (1,000 c.p.m. pmol^{-1}) plus exogenous pKN182 DNA (30 µg ml^{-1}) without or with addition of rifampicin (25 µg ml^{-1}). The reactions were stopped with EDTA and the DNA was purified as described in Fig. 2 legend. Aliquots of the untreated control (lane 1) and of the rifampicin-treated sample (lane 2) were electrophoresed in a 1.0% agarose gel for 3.5 h at 100 V. The gel was stained with ethidium bromide and autoradiographed. The arrows indicate the positions of supercoiled pKN177 and pKN182 DNA, respectively.

~15-fold. The optimal concentration of exogenous plasmid DNA is about 3 nM; higher concentrations of pKN177 DNA are inhibitory. Similar results were obtained with pKN182 DNA (data not shown).

The requirements of the system (supplemented with exogenous pKN177 DNA) are summarized in Table 1. It was found that the magnesium ion concentration is of critical importance. Optimal dTMP incorporation requires 9–11 mM MgCl$_2$, whereas almost no DNA synthesis can be detected at 5 or 15 mM MgCl$_2$. In addition, the system requires not only deoxynucleoside triphosphates but also all four ribonucleoside triphosphates and is strictly dependent on an ATP-regenerating system. A variable degree of stimulation, depending on the extract preparation, was noted on adding cyclic AMP and/or an amino acid mixture. The replicative capacity for mini-R1 DNA is lost if the extract is prepared in the absence of dithiothreitol.

Having optimized the reaction conditions, we investigated whether exogenous pKN177 DNA can also replicate in extracts from *E. coli* cells that do not carry any plasmid. It was found that extracts from plasmid-free cells can catalyse a significant amount of DNA synthesis when supplemented with pKN177 DNA (Table 2). However, the replicative capacity for pKN177 DNA of such extracts is an order of magnitude lower than that of extracts from cells carrying this plasmid. In comparison, exogenous ColE1 DNA is replicated by both types of extracts with similar efficiency.

To characterize the reaction product, the *in vitro* synthesized DNA was analysed by alkaline CsCl gradient centrifugation (Fig. 2). To distinguish between endogenous and exogenous DNA, extracts from C600 (pKN177) were supplemented with pKN182 DNA. Comparison of the two superimposed gradients shown in Fig. 2*a* indicates that the exogenously added pKN182 DNA not only functions as template for its own replication but also stimulates the synthesis of the endogenous pKN177 DNA. This stimulation is R1 specific as it was consistently observed on addition of mini-R1 plasmids but not on addition of ColE1 DNA (data not shown).

The incorporated label is found in two peaks of fast-sedimenting DNA, corresponding to closed-circular monomers of pKN177 and pKN182. Only a minor fraction of the label sediments in the position of open-circular DNA. Similarly, the reaction product of pKN177 DNA replication in extracts from plasmid-free cells consists predominantly of supercoiled monomeric plasmid DNA (Fig. 2*b*). No peak of labelled DNA fragments sediments more slowly than unit-length single strands, as would be expected for material derived from replicative intermediates[4,8]. These results indicate that the *in vitro* system carries out the complete replication of both endogenous and exogenous mini-R1 DNA without detectable accumulation of partially replicated molecules.

The effects of various antibiotics on mini-R1 replication *in vitro* are shown in Table 3. As expected for semi-conservative DNA replication, the incorporation of dTMP is entirely blocked by inhibitors of DNA gyrase (novobiocin and oxolinic acid)[9,10]. An unexpected finding is the complete inhibition of mini-R1 replication by antibiotics which interfere with different stages of protein synthesis (chloramphenicol, puromycin, streptomycin and tetracycline). Plasmid replication evidently is coupled tightly to translation, because DNA synthesis stops within 10 min of the addition of chloramphenicol (data not shown). Plasmid DNA synthesis is also strongly affected by rifampicin, a specific inhibitor of RNA polymerase. The residual rifampicin-resistant incorporation observed with extracts from plasmid-carrying (but not plasmid-free) cells is probably due to the presence of endogenous plasmid-specific RNA.

The inhibitory effect of rifampicin on plasmid replication in an extract containing endogenous pKN177 DNA and exogenous pKN182 DNA was further analysed by agarose gel electrophoresis of the reaction products (Fig. 3). Comparison of the ethidium bromide-stained gel with the autoradiogram shows that the ^{32}P-label is incorporated exclusively into supercoiled plasmid DNA. Interestingly, only the endogenous pKN177 DNA but not the exogenous pKN182 DNA can function as a template for rifampicin-resistant DNA replication.

Table 3 Effects of antibiotics on pKN177 DNA synthesis

Addition (μg ml^{-1})	Source of extract C600 (pKN177)	C600
None	101.7	13.8
Novobiocin (25)	1.6	0.2
Oxolinic acid (50)	2.4	0.3
Chloramphenicol (100)	1.1	0.3
Puromycin (100)	1.5	0.2
Streptomycin (100)	1.6	–
Tetracycline (100)	1.4	—
Rifampicin (25)	14.9	0.1

Standard reaction mixtures (25 μl) were supplemented with exogenous pKN177 DNA (30 μg ml^{-1}) and 0.5 mM of each of the 20 amino acids. ^{3}H-dTTP (500 c.p.m. pmol^{-1}) was used as radioactive label. Incubations were performed in duplicate at 30 °C for 60 min. Values shown are dTMP incorporated (pmol).

The data presented here demonstrate that mini-R1 replication *in vitro* requires both transcription and translation. A requirement for *de novo* RNA and protein synthesis has previously been observed *in vivo* for the uncontrolled replication of a runaway-replication mutant of R1 (ref. 11). It was postulated that the translation product is an R1-specific initiator protein that is not re-usable. Indirect evidence that R1 carries a gene essential for autonomous replication has been presented by Kollek *et al.*[12]. This gene is considered analogous to the *rep A* gene of the closely related R100 plasmid[13]. Evidently, mini-R1 replication *in vitro* is also strictly dependent on the plasmid DNA-directed synthesis of an essential replication protein. The R1-specific stimulation of endogenous plasmid replication by addition of exogenous mini-R1 DNA can thus be explained by assuming that the presumptive *rep A* gene product is a *trans*-acting protein. The transcription requirement would then reflect the synthesis of a messenger RNA for this protein. On the other hand, the differential effect of rifampicin on endogenous and exogenous plasmid DNA synthesis (see Fig. 3) indicates that a *cis*-acting RNA associated with the template DNA may also be involved in R1 DNA replication.

Results similar to those presented here for pKN177 and pKN182 have also been obtained for other mini-R1 plasmids (R.D., unpublished data). The plasmids tested include pKN1562 the copy control of which corresponds to the parent R1 plasmid[14]. Therefore, this *in vitro* system is not restricted to a particular type of Cop mutant and should prove useful for a detailed analysis of R1 replication control.

We thank Dr Søren Molin for supplying the plasmids pKN177 and pKN182 and for helpful suggestions, and Dr Heinz Schuster for providing support and laboratory facilities. R.D. was supported by a short-term fellowship from EMBO and a grant from the Danish MRC.

Received 15 August; accepted 21 November 1980.

1. Kolter, R. & Helinski, D. R. *A. Rev. Genet.* **13,** 355–391 (1979).
2. Tomizawa, J. in *DNA Synthesis: Present and Future* (eds Molineux, I. & Kohiyama, M.) 797–826 (Plenum, New York, 1978).
3. Staudenbauer, W. L. *Curr. Topics Microbiol. Immun.* **83,** 93–156 (1978).
4. Inuzuka, M. & Helinski, D. R. *Biochemistry* **17,** 2567–2573 (1978).
5. Inuzuka, M. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5381–5385 (1978).
6. Molin, S. & Nordström, K. *J. Bact.* **141,** 111–120 (1980).
7. Stougaard, P., Molin, S., Nordström, K. & Hansen, F. G. *Molec. gen. Genet.* (in the press).
8. Staudenbauer, W. L. *Molec. gen. Genet.* **145,** 273–280 (1976).
9. Staudenbauer, W. L. *J. molec. Biol.* **96,** 201–205 (1975).
10. Staudenbauer, W. L. *Eur. J. Biochem.* **62,** 491–497 (1976).
11. Uhlin, B. E. & Nordström, K. *Molec. gen. Genet.* **165,** 167–179 (1978).
12. Kollek, R., Oertel, W. & Goebel, W. *Molec. gen. Genet.* **162,** 51–57 (1978).
13. Yoshikawa, M. *J. Bact.* **118,** 1123–1131 (1974).
14. Molin, S., Stougaard, P., Uhlin, B. E., Gustafsson, P. & Nordström, K. *J. Bact.* **138,** 70–79 (1979).

Nature Vol. 289 22 January 1981

Highly conserved lipoprotein assembly in teleost and amphibian yolk-platelet crystals

R. H. Lange

Institute for Anatomy and Cytobiology, University of Giessen,
Aulweg 123, D-6300 Giessen, FRG

Ohlendorf *et al.*[1] have reported that yolk-platelet crystals in the frog, *Xenopus laevis*, were orthorhombic (space group $P2_122_1$) and consisted of elongated asymmetric units (~$5.5\times11.5\times25$ nm). The presence of an orthorhombic lattice in amphibian yolk platelets fulfilling the symmetry requirements of crystals had been missed in earlier structural interpretations[2,3]. A comparative study of teleost, urodele and anuran yolk-platelet crystals performed in our laboratory has shown the widespread occurrence among equatic vertebrates of an orthorhombic lattice with unit-cell sides similar to those of *Xenopus*. However, we report here that the space group ($P2_12_12_1$) and certain features of the macromolecular complex are at variance with those reported for *Xenopus*[1].

The material studied consisted of the amphibians *Rana temporaria* (gastrula) and *Triturus* sp.[4] (uncleaved, two-cell and blastula stages) and the teleost *Pelvicachromis pulcher*[5] (Cichlidae; immature and mature eggs). Thin-sectioned specimens were fixed (1% glutaraldehyde/1% OsO_4/0.1 M phosphate buffer) and subjected to two very different embedding procedures (epoxy-resin and glutaraldhyde–urea[6]). On applying electron diffraction and tilting, the presence of an orthorhombic lattice was confirmed in all three cases. This was achieved by establishing the presence of symmetry *mmm* for the reciprocal orthorhombic lattice in each case. Moreover, the unit-cell sides *a*, *b*, *c* were almost identical among the three species studied and—falling in the range derived by Ohlendorf *et al.*[1,7] for wet and lyophilized platelets—were also very close to those of *X. laevis* (Fig. 1 legend). This observation supports our view that we are dealing with much the same structure in teleost and amphibian yolk platelets and that our data are also relevant to the findings in *Xenopus*[1,7]. The evolutionarily conservative behaviour of the molecular assembly is underlined by the very similar appearance of 12 indexed diffraction patterns (for [100], [010], [001], see Fig. 1) as well as seven identified crystal projections (electron images) in the three species studied.

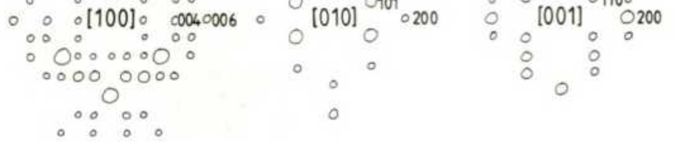


Fig. 1 Electron diffraction patterns (tracings) obtained in *Triturus* sp. when the crystals were oriented as indicated. In the other two species studied, only changes in the intensities of single reflections were noted. The means of unit-cell lengths as determined from such patterns in the three species had the following rangers (in parentheses the X-ray data from *X. laevis* for dried[7] and wet[1] preparations are given for comparison): $a=8.3\ldots9.1$ ($7.9\ldots8.9$) nm; $b=15.8\ldots17.7$ ($15.8\ldots17.2$) nm; $c=16.6\ldots19.3$ ($17.6\ldots19.6$) nm; the lower limit of our data refers to epoxy-resin embedding following dehydration in acetone, the upper limit to glutaraldehyde–urea embedding in the partly hydrated state.

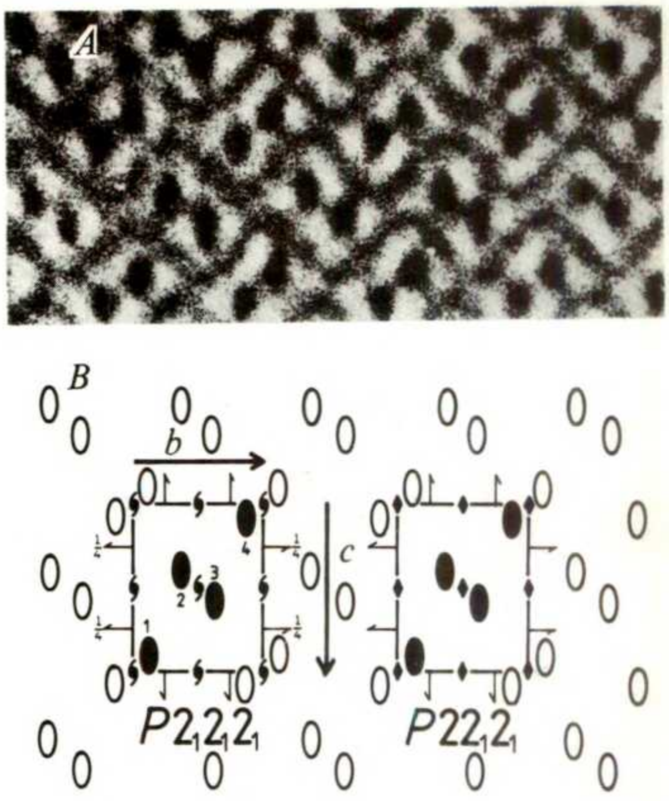


Fig. 2 Electon micrograph of a teleost yolk-platelet crystal viewed along the short axis *a* of the unit cell (*A*) and tracings of the same [100] projection (*B*) made approximately to the same scale. Orientation and size of the unit cell have been unequivocally defined by corresponding diffraction patterns; the cell may, however, be displaced. The footprint-like pattern shown can only be interpreted by assuming 2_1-screw-axes parallel to the unit-cell sides *b* and *c*. Note that these two screw-axes have unique positions in this projection, there are no alternative positions for them. If the primitive orthorhombic lattice has been correctly assigned then only two space group interpretations are possible[11] ($P2_12_12_1$ and $P22_12_1$, but not $P2_122_1$); these interpretations are illustrated in *B*. $P22_12_1$ implies a 2-fold symmetry axis (⧫) perpendicular to the plane of the paper in the middle of each pair of dark elliptic bodies and thus a location of these bodies in the same height along this axis; $P2_12_12_1$ has a 2-fold screw-axis (❟) at this site, implying that the height of neighbouring dense bodies differs by $a/2$, which is in fact shown to be the case (Fig. 3). *A* is a high-power enlargement of a raw electron micrograph originally recorded on 35-mm film at magnification ×25,000 and shows a certain degree of distortion due to sectioning. Dense particles (probably phosvitin complexes, see text) are numbered in the same way as in Fig. 3*B*. International symbols[11]; $b\sim16.5$ nm.

The absence of reflections (100, 300, 010, 030, 050, 001, 003, 005, 007) in the electron diffraction patterns suggests the presence of 2-fold screw-axes parallel to all three unit-cell sides, that is, space group $P2_12_12_1$. Conclusive proof for the space group, however, can be obtained in this case from the mass distribution in electron micrographs (Figs 2, 3). The most revealing projection with respect to molecular arrangement is the previously unstudied projection along the *a* axis ([100], Fig. 2), the second most revealing that along the *c* axis ([001], Fig. 3). It is evident from Figs 2 and 3 that there are four asymmetric

0028–0836/81/040329—02$01.00

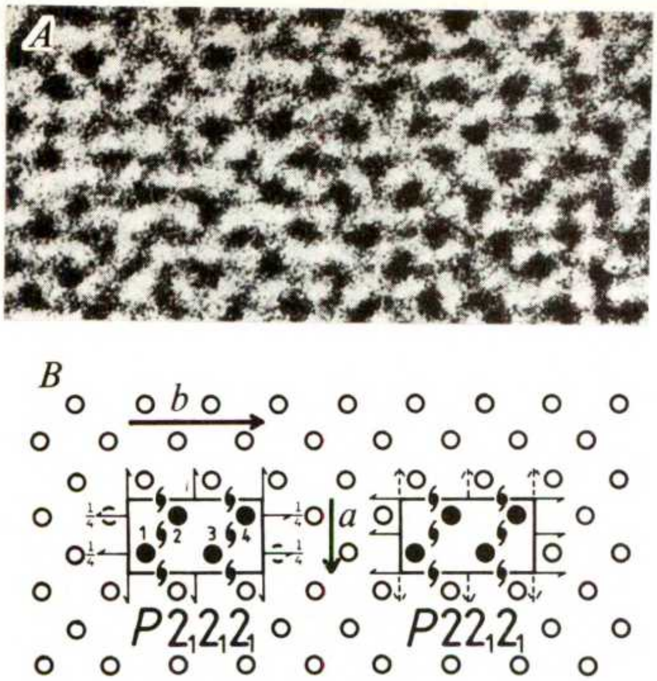


Fig. 3 Electron micrograph of a newt platelet crystal viewed along the *c* axis (*A*) and schematic tracing of the same [001] projection (*B*) made approximately to the same scale. It can be seen that only a screw diad (space group $P2_12_12_1$) and not a rotation diad (space group $P22_12_1$) parallel to unit-cell side *a* correctly describes the matter distribution. Note that positioning of these additional symmetry elements is not deliberate for geometrical reasons[11]. The rotation diad not consistent with the observed mass distribution is shown dashed on the right half of *B*. Compare *A* with the negatively stained preparation in Fig. 1 of the report by Ohlendorf *et al.*[1]. For further details see Fig. 2 legend.

units per unit cell, as in *X. laevis*[1]. There is no evidence that the complex exceeds the length of the *b* axis[1]; however, this point awaits confirmation by further analysis, as does the positioning of a local 2-fold axis[1] of the dimeric structure[8].

To explain the discrepancies between the findings in *X. laevis*[1,7] and our amphibian and teleost data, we should consider the methods used. As shown by investigations using negative staining techniques in electron microscopy of protein molecules[9], there is commonly an overlap of negative and positive staining of such preparations. Such overlap is inevitable and must be considerable when a highly anionic protein like the phosvitin moiety[8] of the yolk-platelet subunit is stained with a cationic dye containing uranyl ions[1]. The dark spots in the positively stained crystals in [100] and [001] projection (Figs 2, 3)—unrecognizable in unstained sections but quite evident in sections stained with uranyl acetate at low *p*H (3.5)—give the most probable location of the phosvitin moiety[2]; in contrast to previous views[2], however, it accommodates both phosvitin molecules of the dimeric complex[8]. This significant part of the complex may largely have escaped representation in the negatively stained preparations of *X. laevis* platelets[1]. In fact, the negatively stained preparation (Fig. 1 of ref. 1) is almost identical to the [001] projection (Fig. 3) in the three species studied here following positive staining of crystal sections with uranyl acetate and lead citrate. In a negatively stained preparation, the space occupied by stain is interpreted as being void of the macromolecule under study; this makes it very unlikely that the results of the three-dimensional reconstruction of the amphibian lipovitellin–phosvitin complex[1] is reliable.

The procedure used in the present study yields results comparable with X-ray data of protein crystals[10], except that it produces a reduction in length of some 10% (due partly to dehydration, especially for epoxy-resin embedding), various degrees of distortion (due to sectioning) and a resolution limit between 1.5 and 2 nm. All our findings were reproduced—and most of our work was indeed done—using glutaraldehyde–urea embedding in the partially hydrated state[6]. The original authors of this procedure have proved that the extremely lipid-rich myelin sheath does not lose its natural birefringence and dimensions when subjected to this embedding method[6]. This implies that the material studied in our laboratory has been preserved as adequately as possible. The sectioning process, furthermore, precludes the bias produced by splitting isolated platelets along preferential planes[1] and thus makes all crystal projections available with the same probability.

I conclude that teleost and amphibian yolk platelets are quite similar assemblies of lipoprotein complexes (orthorhombic, $P2_12_12_1$, almost identical unit-cell lengths for all species studied so far) and, therefore, contain a macromolecular constituent of highly conserved tertiary structure. Space group $P2_122_1$ and the three-dimensional shape of the asymmetric unit as described in isolated yolk platelets of *X. laevis*[1] do not agree with observations made on our comparative material.

I thank Professor S. Haussühl for crystallographic advice, Mr G. Magdowski for technical assistance and the Deutsche Forschungsgemeinschaft for support.

Received 1 July; accepted 20 November 1980.

1. Ohlendorf, D. H., Wrenn, R. F. & Banaszak, L. J. *Nature* **272,** 28–32 (1978).
2. Wallace, R. A. *Biochim. biophys. Acta* **74,** 505–518 (1963).
3. Honjin, R. in *Recent Progress in Electron Microscopy of Cells and Tissues* (eds Yamada, E. *et al.*) 95–108 (Thieme, Stuttgart & Igaku Shoin, Tokyo, 1976).
4. Lange, R. H. *Proc. 7th Eur. Congr. Electron Microscopy*, Vol. 2 (eds Brederoo, P. & de Priester, W.) 624–625 (Leiden, 1980).
5. Lange, R. H. *Cell Tissue Res.* **209,** 511–513 (1980).
6. Peterson, R. G. & Pease, D. C. *J. ultrastruct. Res.* **41,** 115–132 (1972).
7. Ohlendorf, D. H., Collins, M. L. & Banaszak, L. J. *J. molec. Biol.* **99,** 143–151 (1975).
8. Taborsky, G. *Adv. Protein Chem.* **28,** 1–210 (1974).
9. Unwin, P. N. T. *J. molec. Biol.* **98,** 235–242 (1975).
10. Lange, R. H., Blödorn, J., Magdowski, G. & Trampisch, H. J. *J. ultrastruct. Res.* **68,** 81–91 (1979).
11. *International Tables for X-Ray Crystallography* Vol. 1, 104–105 (Kynoch, Birmingham, 1969).

Corrigenda

In the letter 'West Antarctic ice sheet fluctuations in the Antarctic Peninsula area' by D. E. Sugden and C. M. Clapperton, *Nature* **286,** 378–381 (1980), the correction factor of ~750 radiocarbon years for Antarctic Peninsula shells was added to rather than subtracted from the Holocene ^{14}C dates of 6,930 ± 60 and 7,200 ± 50 yr BP for the outer and inner fractions of SRR-1500. The adjusted dates should be 6,180 and 6,450 radiocarbon years BP respectively. The discussion concerning the possibility of an ice-free George VI Sound in the Holocene thus refers to conditions 6,000–6,500 radiocarbon years ago and not ~8,000 years ago. This change does not affect the conclusions of the paper.

In the letter 'Selective killing of mycoplasmas from contaminated mammalian cells in cell cultures' by Marcus *et al.*, *Nature* **285,** 659–661 (1980), the footnote to Table 2 should refer to a 400-W lamp, not 40-W as published.

Erratum

In the letter 'Expression of H-2, lamin and SV40T and TASA on differentiation of transformed murine teratocarcinoma cells' by Knowles *et al.*, *Nature* **288,** 615–618 (1980), two lines were omitted from the bottom of the left-hand column on page 617; it should read:

...... differentiates seems to express the H–2D^b private specificity. It has been observed that somatic cell hybrids between F9 and differentiated cells can ......

BOOK REVIEWS

. . . and when the cupboard is bare?

Kenneth E. Boulding

The Lean Years: Politics in the Age of Scarcity. By R. Barnet. Pp.349. (Simon and Schuster: 1980.) $12.95. To be published in the UK, in April, by Abacus, pbk £2.50. *Energy: From Opulence to Sufficiency.* By J. H. Krenz. Pp.264. (Hemisphere/Holt, Rinehart & Winston: 1980.) $24.95, £15.50.

THESE two books have a somewhat similar theme, but are about as different in style and approach as two books could be. The Krenz volume has a somewhat narrower scope, dealing primarily with energy, whereas that by Barnet takes materials into account; the basic theme of adaptation to the exhaustion of non-renewable resources is common to them both. The book by Krenz is sober, well documented, is based solidly on the best available quantitative information, and presents a rounded picture, even though the author is not afraid to express value judgments on such matters as the future of nuclear energy. The Barnet volume is sparkling, anecdotal in style, rarely presents a complete, well-rounded picture, and is constantly engaged in implied, not too explicit value judgments.

Krenz begins with a clear general consideration of the energy crisis and energy conservation, goes on to a thoughtful discussion of environmental and resource constraints, analyses the situation in regard to fossil fuels, then has a key chapter on more effective energy utilization, which is a better way of describing "conservation". There follows a discussion of "Nuclear Energy: A Faustian Bargain?" which expresses the anxieties and ambiguities involved without dismissing them. Then there is an excellent section on alternative sources of energy, which is both sober and sobering to the over-optimistic. The book concludes with a discussion of the need for a new ethic to deal with the probability of a more austere future.

The Barnet book starts with a rather uneven discussion of the "Post-Petroleum World", and continues with a section on "Guns, Butter and Oil: Changing the Face of Power", which also deals with minerals, food and water, and relates increasing scarcity to what he calls "the new international military order", particularly in the Third World. He concludes with a section entitled "The Global Factory: The Planning of Scarcity", in which he foresees a world employment crisis resulting from the disruption of traditional agriculture and the population explosion. This last chapter ends with a rather vague plea for decentralization without any clear indication as to how to achieve it. Here, Barnet sees little reason for optimism, and what emerges is the general feeling that bad boys will continue to be bad because whoever stops them will be worse. It is, indeed, a gloomy end to liberal hopes.

In terms of suggested solutions to the problem, Krenz, as befits the somewhat narrower scope of his book, lays great stress on the value of conservation and the large technical potential which is latent in conservation practices, in just getting more in terms of human values out of the same input of energy. There are estimates which suggest that we can maintain much the same style of life that we do now with 25 per cent of our current energy input, or with a bit of luck and technical change, maybe 50 per cent. This, in a sense, is a common task which must be shared widely. Barnet, on the other hand, leans towards conspiracy theories of what is the matter with us, does not lay so much stress on conservation, and ends up arguing for a rather vague kind of decentralization as a way of saving us from the concentrated power either of large corporations or of socialist states.

The contrast between the two books illustrates at least one of the virtues of quantified data. The Krenz book, even though it is somewhat abstracted from reality, does give us an overall picture of the proportions of the problem, which the qualitative anecdotal approach of Barnet does not really permit. There is a basic difference, perhaps, between the scientific and the journalistic approach.

Both books would be improved by a little more economics. Although neither of the authors is an economist, Krenz recognizes the importance of economic considerations, though he is perhaps a little over-impressed with purely physical concepts of efficiency. From the human point of view, the only significant concepts of efficiency are those which involve human values. Thus, a process can be highly inefficient in terms of transformation into a particular form of energy, and even highly inefficient according to second law efficiency, that is, transformation of the availability of energy for work. But if the energy output is important by human standards and the input is low-valued, the process can be very efficient in human valuation terms. Thus, the rise of electric power, especially after about 1910, was inefficient by some physical standards but was extremely efficient in terms of human valuations. In commodity output at least, as measured by gross national product, per unit of energy input increased sharply in those decades, largely because of the greater convenience of subdivisibility of electricity as compared to alternative means of transporting energy. Both volumes neglect the problem of energy transport, which is most important from the point of view of human valuations. Energy when and where and in the form in which we want it almost invariably involves its transportation.

Energy: From Opulence to Sufficiency by Krenz can be strongly recommended to those who want an overall survey of the technological problems of energy in fairly non-technical language. The Barnet volume, *The Lean Years,* would appeal to those who enjoy liberal–radical rhetoric, who like to think that everything that is wrong with us is somebody else's fault or the result of a defective political structure of society, which no doubt it often is. Neither book, however, really comes to grips with the way in which decision making and the great network of social interaction actually move us towards overall betterment or worsening. But this, no doubt, is too much to ask. ☐

Kenneth E. Boulding is Distinguished Professor of Economics Emeritus and Director of the Program of Research on General Social and Economic Dynamics at the Institute of Behavioral Science, University of Colorado, Boulder.

Jerrold Krenz — sober and thoughtful

Richard Barnet — sparkling and anecdotal

Progress and prospects in tumour virology

Peter W. J. Rigby

DNA Tumor Viruses. Molecular Biology of Tumor Viruses, Part 2, 2nd Edn. Edited by J. Tooze. Pp.945. (Cold Spring Harbor Laboratory: 1980.) $65, $78 outside USA.

THE first edition of this monograph on the molecular biology of tumour viruses was both timely and influential. It was timely because its publication, in 1973, coincided with the advent of techniques for the physical mapping of viral genomes and it was influential because it acted as an invaluable source for the many molecular biologists, including myself, who began working on tumour viruses in the hope that they would, in conjunction with the new techniques, provide insights into the mechanisms of eukaryotic gene regulation. This second edition thus had much to live up to and it has, in general, succeeded most admirably.

Like its predecessor, the book is a compilation of chapters written by a variety of authors and brought together under the editorship of John Tooze. In his preface Tooze admits to some of the problems that bedevilled the preparation of the book; in particular, it proved impossible to persuade so many authors to produce their material at exactly the same time. While many of these problems have clearly been resolved, there remain cases where it is apparent that closely related chapters were written, or finally revised, at quite different times and there are thus inconsistencies of fact or interpretation. These deficiencies are, however, more than compensated by the book's strengths. It aims successfully at a wide audience, being written in a clear, didactic style quite approachable by final-year undergraduates, yet it is also sufficiently comprehensive and penetrative to be of great value to established workers in the field.

The first section comprises six chapters concerned with the smallest DNA tumour viruses, the papovaviruses and papillomaviruses. Viruses of the former class are now understood in considerable detail which is reflected in separate chapters on genomic organization, the lytic cycle, cellular transformation and genetics. Each of these contributions was written by different authors which leads to some unevenness of approach due principally, one suspects, to the writing having been done at different times. This shows particularly in the treatment of the middle T-antigen of polyoma and of the cellular 53K protein found in SV40-transformed cells. I particularly enjoyed the chapter by Nick Acheson on the lytic cycle of SV40 and polyoma. This deals in some detail with experimental problems and gives the reader an excellent feel for the certainties and uncertainties of the area. The chapters on human papovaviruses and on papillomaviruses are quite different in content, being necessarily much more biological in approach. However, the molecular biology of the human papovavirus BK is dealt with poorly; its DNA sequence, which appears in one of the appendices, is hardly discussed.

Adenoviruses are covered in three chapters, two written by Jane Flint alone, the other by her in collaboration with Tom Broker. That one author was involved in all three chapters shows clearly in a much more uniform approach and consistent treatment. The chapter on lytic infection is truly excellent, being both comprehensive and detailed yet clearly written and easy to follow. It also deals superbly with some of the less-known regions of tumour virus research, including an outstanding discussion of adenovirus assembly.

The chapter on herpes simplex viruses, by Patricia Spear and Bernard Roizman, is by far the largest in the book and is also the least satisfactory. It would have been much more appropriate for a specialized review. In many places the textual discussions would have benefited from the provision of appropriate diagrams while, conversely, it is the only chapter to use extensive figures of primary experimental data. These are discussed in a style so condensed that it will be impenetrable to all but experts. The chapter is also somewhat unbalanced. It deals too superficially with herpes genetics and is too heavily biased towards the work and opinions of the authors.

The final pages are occupied by six appendices giving the DNA sequences of SV40, polyoma and BK, and of the transforming segments of adenovirus types 5, 7 and 12. Also presented are lists of restriction sites, maps of mRNAs and putative protein sequences. This section represents an excellent innovation and will be of the greatest value. I have only one criticism of this section, albeit a severe one. The sequence of SV40 has recently been modified by the addition of an extra, small restriction fragment. While this is referred to in the introduction to the appendix, the sequence used does not include it. It would have been worth correcting this, even at the price of a slight delay in publication.

This book is as timely as its predecessor, detailing the results of one phase of the work and setting the scene for the studies that are to come. It will, therefore, be as influential. I fully share the editor's regret that the increased size of the book has necessitated a significant increase in price and sincerely hope that a paperback edition will be forthcoming so as to put the book within the reach of the many students who ought to own a copy. Economics notwithstanding, this book is mandatory reading for those wishing to obtain a general view of tumour virology, for those wishing to become involved in this most exciting area of research and for those currently involved in it. The editor and the authors are to be congratulated on undertaking so well what must have seemed to them an almost impossible task. I can only hope, on behalf of all in the field, that in five years time they will have the courage to attempt a third edition. □

Peter W. J. Rigby is a member of the Cancer Research Campaign Eukaryotic Molecular Genetics Research Group and a Lecturer in Biochemistry at the Imperial College of Science and Technology, University of London.

Just beyond the pale

Paul Davies

Gravity, Particles, and Astrophysics. By Paul S. Wesson. Pp.188. (Reidel: 1980.) Dfl.65, $34.

COSMOLOGY and astrophysics attract an enormous number of speculative theories, not all proposed by cranks, but many of them written somewhat tongue-in-cheek. Paul Wesson is a man with an encyclopaedic knowledge of those grey areas somewhere between the canonical models of orthodoxy, and the manifestly absurd, that manage to slip into the reputable journals. His previous book, *Cosmology and Geophysics* (Adam Hilger, 1978), dealt lucidly with a variety of "variable G" theories, in which the Newtonian gravitational G is assumed to vary with time. It is a theme continued and updated in this latest work, but accompanied by a range of other, similar, conjectures: particle creation, anomalous redshifts, tired light and cosmic numerology, for example.

The author's subjects range ambitiously from the planets to the primaeval furnace. Scarcely an unorthodox idea escapes attention somewhere. The professional cosmologist will be alternately delighted and dismayed to see many half-forgotten theories receiving an airing. From the Lyttleton–Bondi cosmology to modern gauge theories of gravity, Wesson gets around to them all in some context or other.

Like it or not, many of these unusual theories have been extensively, even infuriatingly, discussed in the scientific literature for years. Some of them, such as the so-called big number coincidences or the varying *G* theories, have occupied the attentions of eminent physicists and astronomers. One should not blame the author for attempting a comprehensive presentation of such issues just because they command only minority support.

My real criticism of this book is that it reads rather like a shopping list. In his enthusiasm for including everything, the author has only managed to achieve a compendium of speculations, and not a coherent discussion. References are so liberally used that the text becomes impossible to read in some parts. On one page I counted one hundred such interjections. This is fine in a reference source for non-standard cosmology and astrophysics, but will not satisfy the demand by conservative "straight thinkers" to know something of the work of their more adventurous-minded colleagues.

Of course, a subject does not become respectable merely by quoting a vast array of big names. Even Einstein patronized some odd theories. Nor should one be persuaded that all unconventional ideas are in some mysterious way consistent with each other. This book is written in a style that leads one to believe there is a sort of grand, coherent theory of unconventional physics and cosmology existing alongside the more publicized one. It gives the impression of two sciences: the one to which most professionals subscribe, and another, almost the same in its predictions, but with subtle differences. Wesson is fair in drawing attention to some of the shortcomings of the latter, but clearly believes that his several hundred "grey area" references are probing closer to the truth than many of us would be prepared to acknowledge. □

Paul Davies is Professor of Theoretical Physics at the University of Newcastle upon Tyne.

Agricultural prognoses

Kenneth Blaxter

Britain's Future in Farming. Edited by Sir Frank Engledow and Leonard Amey. Pp.164. (Geographical Publications: Berkhamsted, Herts, UK, 1980.) £9.

THE farming industry accounts for less than 3% of the labour force in the United Kingdom and contributes about the same percentage to its gross domestic product. Yet it is the most obvious of our industries since it occupies the major part of our countryside and produces most of the food we consume. The industry has much changed with time; it is now capital intensive and heavily dependent in some of its sectors on sophisticated inputs from other parts of the economy. These changes are relatively recent and, as the authors of this book state, "a highly urbanized economy with a romantic view of the countryside has yet to come to terms with the unpicturesque aspects of an intensive and efficient agriculture". Questions necessarily arise about the future of farming, in turn inevitably implying questions about the policy to be adopted and the goals to be sought; these considerations are basic to this volume.

The book is the result of four years of study by a group of 15 people knowledgeable about farming, food provision, forestry, agricultural science, banking, and the social and political history of land use. None of the chapters is assigned to an author, but the discerning can detect where initial responsibility rested for written styles are revealing. There is thus a collective responsibility and this at times seems to have engendered some collective irresponsibility. Surely some of the statements made cannot have been agreed by all. Is it reasonable to suppose that all agreed "that the Ministry of Agriculture [MAFF] should resume its role in education and research and subsume the Agricultural Research Council [ARC]"? Or that security of tenure for farmers should be withdrawn forthwith? Such statements in the text do not appear in the conclusions with which no doubt all agreed. The first becomes transformed to a plea for better integration and a suggestion that attention should be drawn to the relationship of MAFF and ARC, while the second does not appear at all.

A few provocative statements, however, do not detract from the overall thesis and, on the contrary, perhaps enhance what is a sensible account of the problems related to the management of our resources of land. It should do much to show that the interests of agricultural and urban communities in reality have much in common and that certain broad principles related to land use may well be sufficient to provide a basis on which to develop policies to meet new eventualities. The conclusions and recommendations are admittedly detailed — there are 28 of them — but the whole tenor of the book relates to the theme that the source of renewable land must be preserved, that the pressures upon it whether direct or indirect must be reduced and farming practices themselves should be designed to provide a flexibility to accommodate inevitable and continued change. □

Sir Kenneth Blaxter is Director of the Rowett Research Institute, Aberdeen.

Practical vacua

John Yarwood

A User's Guide to Vacuum Technology. By John F. O'Hanlon. Pp.402. (Wiley-Interscience: 1980.) £13.35, $31.20.

JOHN O'Hanlon is a member of the research staff at IBM in the USA. As a result this book is really up-to-date. He has written the most useful and thoughtful account of vacuum technology which has appeared in recent years, and one which will be of value to the student, technician, engineer, scientist and technical manager. As would be expected from an author with such a background, the book (according to his preface) emphasizes "the operation and selection of equipment for processes used in semiconductor, optics, and related technologies". But he does himself an injustice — this book is much wider in scope than such a statement would appear to indicate.

SI units are used throughout, with the exception that pumping speed (now called "volume rate of flow" — not a good term although scientifically accurate) is given in l/s or m³/h, depending on the type of pump, and the diameters of the top flanges of diffusion pumps are given in inches. The section on essential gas theory covers gas properties, kinetic theory, basic relevant formulae (without proofs), gas flow and the measurement of pumping speed (where the work of the International Standards Organization has not been acknowledged; all the credit seems to go to the USA). Succeeding chapters are up-to-date, apposite accounts of real value, dealing with a wide range of equipment — gauges, pumps, analysers — and associated topics.

The book is rounded off with six appendices and, overall, can be recommended as a most useful and comprehensive volume.

John Yarwood was formerly Head of Physics at the Polytechnic of Central London, and is the editor of Vacuum.

UV photobiology

Philip C. Hanawalt

Biological Effects of Ultraviolet Radiation. By Walter Harm. Pp.216. (Cambridge University Press: 1980.) Hbk £15, $29.95; pbk £4.95, $9.95.

WE HAVE long appreciated the essential role that light plays in the maintenance of life forms on our planet through the processes of photosynthesis. In recent years, however, we have become increasingly aware of the deleterious effects of the shorter wavelengths of light in the Sun's spectrum. We have learned that ultraviolet light (UV) damages the precious DNA, resulting in mutations, the death of cells and sometimes malignant transformation. The latter effect would be intolerably frequent were it not for cellular enzymes that constantly monitor the DNA to identify and repair the lesions that UV inflicts. This is dramatically illustrated by the rare hereditary disease, xeroderma pigmentosum, in which DNA repair deficiency renders the victims exquisitely sensitive to sunlight-induced skin tumours.

Biological effects of UV was an excellent choice of topic for this first volume in a new series sponsored by the International Union of Pure and Applied Biophysics, since UV is both an important biophysical probe for DNA repair systems and an environmental hazard of some concern. The books in this series are meant to serve as specialized texts for advanced students in particular fields of biophysics. The author, Walter Harm, was one of the pioneers in the field of UV photobiology and he was therefore in a good position to provide the kind of historical perspective that could bring the subject to life and excite students. Unfortunately this monograph doesn't really gain much vitality as it plods ponderously through the details of classical bacterial UV photobiology.

Without first surveying the range of interesting biological effects, the first chapter launches into electronic excitation and the practical nuts-and-bolts details of experimental UV photobiology. Chapter 2 carries on with more practical advice for the researcher, including some useful numerical examples to illustrate important points. A carefully detailed consideration of UV interactions with biologically important molecules is followed by a not-so-useful chapter on classical target theory and survival curve shapes. (The word "multihitlike" was a new one to me.) I would have preferred more discussion of why shoulders on survival curves may be explained by existence of repair systems. The terms excision-repair and post-replication repair are introduced but not defined until three chapters later.

The subject of enzymatic photoreactivation is treated in great detail, reflecting the recent research interests of the author. In contrast, the discussion of dark repair systems is incomplete and contains a number of inaccuracies, such as a reference to sequential action of the *uvrA* and *recA* systems (p.107). The book is nearly devoid of discussion of our current understanding of the enzymology of dark repair and its genetic control. Yet this is perhaps one of the most exciting areas of research in UV photobiology.

The predominant emphasis in the text is on UV inactivation of bacteria and their phage, even in the chapter on solar UV radiation effects. Sunlight effects on human skin are relegated to less than two pages followed by a brief chapter on UV carcinogenesis and an even shorter one on sensitized UV effects on DNA.

I must say that I would find it difficult to teach a course using this book as a text but I find it quite useful as a reference work. It does provide a comprehensive treatment of the now classical studies on UV effects in prokaryotes. It might have been more aptly titled *Microbial Effects of UV Radiation*; as Dr Harm emphasizes, much of our understanding in UV photobiology and DNA repair has been based upon the solid foundation of early studies with simple prokaryotes.

Philip C. Hanawalt is Professor of Biology and Dermatology at Stanford University, and Director of the Graduate Program in Biophysics.

Social science and the stomach

John Yudkin

Consuming Passions: The Anthropology of Eating. By Peter Farb and George Armelagos. Pp.279. (Houghton Mifflin: 1980.) $12.95.

THE past 25 years have seen not only a rapid growth in research in nutrition, but the increasing recognition that nutrition is as much a social as an experimental science. Its wide span is encompassed in the short definition of nutrition as the study of the relationship between people and their food. Thus, there are now fewer academics who consider nutrition to be a section of biochemistry or of physiology, and few among the growing number of university departments of nutrition that do not include in their courses relevant aspects of social nutrition from economics, sociology, anthropology and psychology.

This book will certainly be of interest to nutritionists and dietitians, as well as to anthropologists, but its appeal is very much more general since it contains a mass of fascinating information. Its wide range is best indicated from this quotation near the end of the work: "The thesis of this book has been that to know what, where, how, when and with whom people eat is to know the character of their society". For example, you will learn from the book that "cannibal" is derived from "Carib" (as in "Caribbean") and that a "companion" is the person with whom you eat bread; that among the people of Trobriand it is unseemly for couples to eat together before marriage, just as it is (was?) for unmarried Western couples to sleep together; that cultural customs such as not eating pork or the reverence of the Hindu for cattle often have a biological or ecological basis.

The authors deal first with the evolution of human dietary patterns, how they have adapted to the availability of local foods, and the differences called for, and met, at different phases of the life of the individual. There follows a description and suggested explanation for the myths and religious concepts that influence eating behaviour, and the rules that different people have adopted towards food in their relationships with family, clan and strangers. There is considerable discussion of food preferences and avoidances, and the development of national cuisines.

It would be pleasing to be able to say that this book had achieved a synthesis between those aspects of eating that belong to the social sciences and those that belong to the natural sciences. The fact is, however, that until recently at any rate most people who have become professionally interested in nutrition have come through the separated and narrowly confined and demarcated pathways of these two groups of disciplines. It is understandable then — but a great pity — that the data on the nutrient content of foods have not been properly checked.

It is not true that one person may "burn calories ten times faster" than another "of the same age, sex and physical proportions". It is misleading, too, to worry about the fact that a portion of canned peas contains 100 times as much salt as do fresh peas, when in absolute terms this would still be about one-fiftieth of the amount taken in a day; more importantly, it ignores the fact that people add very much more salt when cooking vegetables than that present in the canned peas. The plausible suggestion that pica usually has a nutritional significance — for example, that eating clay provides minerals lacking in the diet — has been repeatedly disproven.

One last grumble is that the authors have adopted the somewhat hysterical current concern about the safety of modern food processing. For example, they say categorically that food colours are carcinogenic, a statement that, of all countries, is highly unlikely to be true in America with its rigid food laws that are especially sensitive about suspected carcinogenicity of food additives.

Nevertheless, for the general reader this is a book that is well and clearly written (except for the use of "human" as a noun!) and is full of interesting facts about what and why we eat what we do. For the specialist reader, too, there is a host of references. If, therefore, you are dubious about the theories of Claude Levi-Strauss on attitudes and behaviour towards food, you will be directed to the writings of several other authors, some of whom agree and some of whom disagree with his views. □

John Yudkin is Emeritus Professor of Nutrition at London University.

CORRESPONDENCE

Continued from page 218

provisions of the First Amendment of the Constitution by permitting the teaching of evolution. This, the suit says, disparages the religion of creationists and in itself amounts to the teaching of a religion — the religion of secular humanism.

The charges in the suit focus on the *Science Framework for California Public Schools, Kindergarten and Grades One through Twelve*, published by the California State Department of Education in 1978. Dr Richard Lemmon and I served on the committee that wrote this publication, a task that took about three years. It involved a steady battle with creationists who resisted the inclusion of statements concerning evolution. On some occasions, we obtained their consent, only to have them change their minds. For example, a statement that "Darwinian evolution is a cornerstone of modern biology" was first permitted, then deleted.

The final document says:

> Living organisms have universal properties, including derivation of energy from outside sources, and reproduction. Evolutionary studies indicate that these organisms are naturally selected from generation to generation, producing descendants with different characteristics and producing variability among populations of living species. The process has been going on so long that it has produced all the groups and kinds of plants and animals now living as well as others that have become extinct.
>
> . . . In addition to reproduction, another characteristic of life is change in its genetic material with passage of time. This process, termed evolution, takes place through changes in DNA. Changes in DNA molecules are produced by mutation, which includes replacement of some DNA bases by others, and recombination in which segments of DNA are added to or subtracted from genes. Duplication of genes sometimes also occurs. Most mutations are harmful and do not persist; they are eliminated by natural selection.
>
> Beneficial mutations occasionally take place and are responsible for the appearance of new characteristics. A third class is "neutral" or "near neutral" mutations.

The Creation Research Society (through Creation-Life Publishers) does not confine itself to matters of religion but also addresses scientific questions. A complaint by a parent in Livermore is that the publication *Dry Bones*, published by Creation-Life in 1979, is being used in lower grades of an elementary school. It contains a dialogue between children and their father, including:

> Q: "Did dinosaurs live when Noah lived, Dad?"
> A: "That's what I think. The way it looks, dinosaurs were drowned in the Flood, too."
> Q: "Did Noah have to take them on the Ark?"
> A: "He surely did". . .
> Q: "I guess[Darwin] thought fossils were proof for evolution."
> A: "No, Darwin said that fossils were 'perhaps the most obvious and serious objection' to his theory."
> A: ". . . Did your friend also tell you about scientists at Oak Ridge National Laboratories?"
> Q: "No, what about them, Dad?"
> A: "They used uranium dating on wood in rocks of the dinosaur group, and got ages of only thousands of years."

THOMAS H. JUKES

University of California, Berkeley

Genes and race

SIR — Some time ago you published a letter from me concerning the use made by the National Front of apparently "scientific" justification of racist arguments. I asked, through your columns, Professors Eysenck and Jensen publicly to dissociate themselves from the use made of their names in apparent support of these views.

I now wish to bring an analogous issue to the attention of readers of *Nature*. Ever since the publication in 1975 of E.O. Wilson's *Sociobiology: The New Synthesis*, the danger has been clear that in due course racists would deploy sociobiology in support of their views. The "New Right" in France has been doing so for some time, and now the breakaway journal from the National Front in Britain, *New Nation*, has done so too. Its first two issues, dated summer and autumn 1980, carry articles entitled: "Nationalism, racialism: products of our selfish genes" and "Science is championing our creed of Social Nationalism", authors respectively John Thornton Bannerman and Richard Verrall (the latter is also the author of a pamphlet entitled "Did Six Million Really Die?").

The first cited of these comes complete with a picture of "Richard Dawkins, author of *The Selfish Gene*", and cites him, Wilson, Maynard Smith and one Travers (presumably Trivers) amongst others. The final section of this article appears alongside an advertisement from "Nationalist Books" which includes such choice titles as ***Debunking the Genocide Myth*** and ***Anne Franck's diary — a hoax.*** Opponents of sociobiology, it is therefore scarcely surprising to note, are categorized as "Marxist" and "Jewish".

The concluding paragraph reads as follows:

> For us, as racial nationalists, this is an important vindication of our position. For it is increasingly clear that it isn't just "bad luck" that our genes don't permit us to live in a Marxist–Rousseauesque egalitarian communist utopian World State of universal altruism. It was an inevitable result of the way evolution works that our genes would not permit us so to live. What the evolutionary theoreticians have shown us is that, with the system of genetic inheritance shared by all vertebrates, the only type of social organization which can evolve, let alone work, is one based upon kinship, upon the ties of blood and of race. Nationalism is not only an integral part of our genetic inheritance, it is an inevitable end product of the evolutionary processes which shaped that inheritance.

May I suggest that it would be in the public interest that John Maynard Smith and Richard Dawkins should clearly dissociate themselves from the use of their names in support of this neo-Nazi balderdash.

STEVEN ROSE

The Open University,
Milton Keynes, UK

Energetic consensus

SIR — It was disheartening to learn from your correspondent that the First European Bioenergetics Conference had shown a "general acceptance of the chemiosmotic concept of energy transduction" and that interest was now moving on from this assured base to the finer molecular details of the problem (*Nature* 4 December 1980, p.432). If there is such consensus on this matter there is very good reason to believe that it could be mistaken.

The chemiosmotic hypothesis is founded upon a very simple and invalidating illogicality: no evidence has ever been given that protons move into the bulk-phase during synthesis. Oxygen-pulse studies which supposedly support the hypothesis are carried out under entirely different experimental circumstances; and to have used the results of such studies to generalize upon the nature of proton movements in ATP synthesis was to introduce an inferential *non sequitur* into the energy transduction argument from the outset. The basic chemiosmotic equation:

$$\triangle p = \triangle\ \psi - Z \triangle p\mathrm{H} \qquad (1)$$

has no meaning if protons are not translocated to the bulk-phase.

We have attempted to state this view since 1974[1] and the accumulating experimental evidence we have obtained has led to the ϱ-zone interpretation of mitochondrial energy transduction in which the proton-motive force has been regarded as a localized and surface electrical force, with a minimal osmotic involvement:

$$\triangle p^{\varrho} \approx \triangle \psi^{\varrho} \qquad (2)$$

Recent experiments showing the generation of negative fixed charges on the inner membrane specifically associated with the energization process[2] have fulfilled the main prediction of the ϱ-zone interpretation and given a further compelling reason to suspect the adequacy of the chemiosmotic view; they also call into question our own formulation of the proton-motive force given in equation (2).

It now seems probable that during synthesis the proton-motive force should be regarded as a coulombic rather than a chemiosmotic force, and that its assessment should be made in terms of charge–charge relationships taking place in a surface region of variable dielectric. Some of the more recondite thermodynamic implications of such a system have already been anticipated by Ashcroft and Coster[3]; meanwhile an unsophisticated estimate of the intercharge distance, l, may be gained from the relationship:

$$W = \frac{q_1 q_2}{\varepsilon l} = 15\ \mathrm{kcal}\ (\triangle Gp)$$

and the speculative use of a now dubious chemiosmotic H^+/ATP stoichiometry of 3. Substituting the dielectric constants for the "limit" media ($\varepsilon_{aq} = 75$; $\varepsilon_{lipid} = 2$) intercharge distances of 0.9 and 33Å are obtained; the functional value may reasonably be expected to lie between these extremes and to be in accord with the requirements of the ϱ-zone interpretation.

F.H. MALPRESS

Department of Biochemistry,
The Queen's University of Belfast, UK

1. Archbold, G.P.R. *et al. Biochem. Soc. Trans.* 2, 751-754 (1974).
2. Archbold, G.P.R. *et al. Biochem, Internat.* 1, (in the press).
3. Ashcroft, R.G. & Coster, H.G.L. *Bioelectrochem. Bioenerg.* 5, 37-42 (1978).

ANNOUNCEMENTS

Awards

Drs J.-P. Mach and S. Carrel, (Lausanne Branch of the Ludwig Institute for Cancer Research) have received the 1980 Cancer Prize of the Swiss Cancer League.

The Charles Frederick Chandler Medal, presented by Columbia University for achievement in pure or applied chemistry has been awarded to **Frank H. Westheimer** (Morris Loeb Professor of Chemistry at Harvard University).

The NASA Group Achievement Award has been presented to **Dr Sam Gulkis** (JPL); **Dr David Jauncey** and **Dr Michael Batty** (CSIRO) for the development of a radio astrometric interferometer in the Southern Hemisphere.

The Council of The Institute of Physics has made the following awards for 1981; Bragg Medal and Prize to **G.E. Foxcroft** (Rugby School); Charles Chree Medal and Prize to **Dr K.A. Browning** (Meteorological Office Radar Research Laboratory, RSRE Malvern); Charles Vernon Boys Prize to **Dr P.N. Pusey** (Royal Signals and Radar Establishment, Malvern); Duddell Medal and Prize to **Dr B.A. Joyce** (Philips Research Laboratories, Redhill); Glazebrook Medal and Prize to **Dr G.H. Stafford** (Rutherford and Appleton Laboratories, Chilton); Guthrie Medal and Prize to **Dr J.C. Ward** (Physics Department, Macquarie University, Australia); Maxwell Medal and Prize to **Dr J.M. Kosterlitz** (Department of Mathematical Physics, Birmingham University); Thomas Young Medal and Prize to **N.J. Phillips** (Department of Physics, Loughborough University of Technology); Prize for the 1980 Graduateship Examination of the Institute to **James Cooke** (Dublin College of Technology).

The Council of the Royal Society invite nominations for the 1981 Royal Society Esso Energy Award which should be sent by 6 February 1981. The Award is given for outstanding contributions to the advance of science or engineering or technology leading to the more efficient mobilzation, use and conservation of energy resources, and consists of a gold medal and a prize of £1,000. Further information from: Dr R.W.J. Keay, The Royal Society, 6 Carlton House Terrace, London SW1, UK.

A fellowship in human anatomy, to be called the Wellcome-Franks Medical Fellowship, has been endowed at St Hugh's College, Oxford, by The Wellcome Foundation Limited. Enquiries to: Ron Champion, The Wellcome Foundation Ltd, 183 Euston Rd, London NW1, UK.

Environmental researchers and scientists — either from university or industry — are invited to compete for Chemviron's fifth 'Award' a $10,000 cash prize to be given in 1982. It is awarded for outstanding work in the area of physical–chemical treatment of potable and waste water, or of sludge resulting from treatment of such waters. Copies of the regulations can be obtained from Dr J. Melbourne, c/o Chemviron, Chaussee de Waterloo 1135, B-1180, Brussels, Belgium. Only entries originating in Europe will be accepted.

The Roussell Prize to be awarded in 1982, is given every two years to a chemist or a biochemist whose work on steroids in therapeutic medicine has been chosen as the most outstanding by an International Committee. Further information from: Professor J. Mathieu, Centre de Recherches, Roussel Uclaf, 102 route de Noisy, 93230 Romainville, France.

The Trustees of the Lady Tata Memorial Trust invite applications for awards for personal support for research on leukaemia, in the Academic Year beginning October, 1981. Candidates with programmes of research on any aspect of malignant disease which may throw light on problems of leukaemia will be eligible for consideration. Further information from: Secretary of the (European) Scientific Advisory Committee, Lady Tata Memorial Trust, MRC Leukaemia Unit, Royal Postgraduate Medical School, Du Cane Road, London W12, UK.

The Paul-Martini-Prize, was initiated by the Medizinisch Pharmazeutische Studiengesellschaft e.V. (Mainz) in association with the Deutsche Gesellschaft Für Medizinische Dokumentation, Informatik und Statistik e.V. to promote the development of scientific methods for evaluation in clinical pharmacological and therapeutical fields. Papers should be submitted by 28 April 1981 either in German or English with six copies to the Medizinisch Pharmazeutische Studiengesellschaft e.V. Bilhildisstr. 2, D 6500 Mainz.

The International Union Against Cancer, with funds provided by the Cancer Research Campaign (UK), will award fellowships for research on cancer to persons on the staff of universities, teaching hospitals, research laboratories or similar institutions. These enable investigators to work abroad to gain new experience in clinical or basic research laboratories or similar institutions. Application forms and additional information from: International Union Against Cancer, rue du Conseil-Général, 3, 1205 Geneva, Switzerland.

Meetings

9-10 February, **New Concepts in Drug and Nutrition Delivery Systems,** Paris (Roberts S. First, Inc. Avenue Marnix 19A, 1050 Brussels, Belgium).

9-13 February **Fundamentals of Microprocessors,** London (Bleasdale Computer Systems Ltd, Francis House, Francis St, London SW1, UK).

9-13 February, **Protein Structure and Function,** Lorne, Australia (Dr R.J. Simpson, St Vincent's School of Medical Research, Victoria Parade, Fitzroy, Victoria, Australia).

12 February, **4th Interdisciplinary Cancer Research Workshop,** New Orleans (Dr P. Politzer, Chemistry Dept, University of New Orleans, Louisianna).

15-17 February, **Proprioception, Posture and Emotion,** New South Wales (Dr D. Garlick, Committee in Postgraduate Education, PO Box 1, Kensington, New South Wales, Australia 2033).

8-15 February, **Immunoglobulin Idiotypes and their Expression,** Salt Lake City (ICN-UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

15-20 February, **Differentiation and Function of Hematopeitc Cell Surfaces,** Keystone (ICN-UCLA, Molecular Biology Institute, University of California, Los Angeles, California 90024).

20 February, **Ulcer Therapy,** Frankfurt (GDCh, Postfach 90440, D-6000 Frankfurt/M8, West Germany).

25-27 February, **Recombinant DNA Research,** San Francisco (Edward Ruffig, c/o Scherago Associates Inc, 1515 Broadway, New York 10036).

22 February — 1 March, **Mechanisms of Chemical Carcinogenesis,** Keystone (ICN-UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

1-8 March **Cellular Recognition,** Keystone (ICN-UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

4-7 March, **Chemical Indices and Mechanisms of Organ-directed Toxicology,** Barcelona (Secretariat, Toxicology 81, 142 Oxford Rd, Cowley, Oxford, UK).

8-21 March, **Basic Concepts and Techniques in Research on Biological Membranes,** Ibadan (Dr Enitan A. Bababunmi, Dept of Biochemistry, University of Ibadan, Nigeria).

9-11 March, **Drug Delivery System Development,** London (A.S. Goldberg, Powder Advisory Centre, PO Box 78 London NW11, UK).

10-13 March, **POWTECH '81.** Birmingham (Clive H.R. Hopkins, Green Dragon House, 64-70 High St, Croydon, Surrey, UK).

nature

29 January 1981

Madrid, Helsinki and exchanges

The reopening of the Madrid conference on the Helsinki agreements will bring very little joy to those directly concerned and little benefit to the rest of us. The proceedings so far have been one long wrangle, a measure of the extent to which East–West relations have deteriorated even since 1973, when the Final Act was signed at Helsinki. The weeks ahead are to be given over to proposals for the further elaboration of the Helsinki agreements (see page 343). For the reasons that have made the conference so far an occasion for the restatement of well-known positions, the second period is likely to see only a rehearsal of familiar opinions, known to be unacceptable to one or other of the states involved. Not much has changed since the proceedings at Madrid began. Sakharov is still in Gor'kii, for example. The replacement of President Carter by President Reagan will matter in the long run, but for the time being is likely to diminish American interest in the proceedings. The best that can be hoped for now is that the remainder of this conference will pass off without too much further damage being done.

One of the potential casualties is that provision of the Helsinki accords under which the Madrid conference itself is being held. Like many other international agreements (the Helsinki Final Act is not a treaty in the formal sense), signatories are invited to meet every so often to consider how things are working out. By comparison with Madrid, the meeting in Belgrade three years ago was more or less constructive. It accomplished little, but it did at least agree that there should be another meeting. The time is approaching in the proceedings at Madrid when people will be asking themselves whether they can properly wish on their successors in office another bout of frustration like the present. It will be unfortunate if they decide too quickly that that question should be left for decision through the normal diplomatic channels. That would assure that nothing like Madrid will happen again. That would be a misfortune.

Among international agreements, the Helsinki accords are a peculiar species. They correspond not to formal contracts among sovereign states to behave in specified ways but, rather, to the kinds of letters of intent that are signed by commercial companies which are persuaded that they have the basis for a deal but which have agreed that they may back off in certain unspecified circumstances. The Helsinki accords thus lie somewhere in between the promises to abstain from alcohol subscribed to by members of temperance movements throughout the world and the fine print on the back of air travel tickets which binds the airline which collects the money to convey the would-be traveller to his destination provided that it is able to do so. While the Helsinki agreements were being negotiated, scepticism about their potential value abounded. In many respects, they have not worked — people's freedom to move elsewhere, for example, is still constrained. On the other hand, the remarkable freedom of journalists to visit Poland in the past few turbulent months may owe something to Helsinki. And advance notification of military manoeuvres may have done something to calm anxiety in Central Europe in recent years. The vagueness of the agreements has also, as predicted, made scope for misunderstanding but also injustice — such as the repression of the Helsinki monitoring groups in the Soviet Union. But the accords exist and, like other international agreements, could be abandoned only by further souring East–West relations.

For the scientific community, what matters most is what the accords have to say about cooperation and exchanges of people. The first flurry of excitement after Helsinki has now been overtaken by events, the banishment of Sakharov to Gor'kii only one of the more obvious. To people in the West, it is intolerable but also mystifying that scientists should be treated like this. In the East, imprisonment, exile and repression are instruments of government entirely consistent with (and thus justifiable by) the old Tsarist feudal principles that a person belongs ultimately to the state. There is nothing in the Helsinki accords to ensure that the Soviet government will behave differently towards its citizens, but there are opportunities to demonstrate that, if it did, important benefits would accrue. This is why flat declarations such as that attributed to Dr Philip Handler during the first frustrating session at Madrid can serve very little purpose. What is needed, and what Western governments should be proposing in the weeks ahead, is a much more imaginative programme of scientific exchanges than those previously arranged. Case law (and common experience) suggests that busy people do not wish to visit unfamiliar places for months on end. Why not, then, shorter visits? And why should not the governments at Madrid commit themselves to the support of participation in run-of-the-mill scientific meetings that would bring in more people from Eastern Europe, not just the Soviet Union?

Naming names in British universities

The crunch is coming for British universities. Indeed, it may have come already. The British government's decision in December that a further £30 million should be taken from the budget of the University Grants Committee in the financial year beginning in April could easily precipitate the crisis that has been looming for the past decade. For although £30 million is a mere 3.5 per cent of the budget as a whole, it will be the first time in recent memory that the universities collectively have needed to operate at less than full capacity. After three decades in which student numbers have steadily increased, universities individually are unaccustomed to turning away students for whom there is space. Worse still, the university system is not equipped to decide within itself how the consequences of even a modest reduction of business should be shared. Must all universities decline by more or less the same proportion? Or is the misery to be concentrated in a few places and, if so, which?

Inevitably, confusion abounds. The Secretary of State for Education and Science, Mr Mark Carlisle, said on 6 January that he saw no reason why the latest budget cut should prevent universities from keeping up student numbers. The University Grants Committee, on the other hand, has issued a dark warning to universities that if they seek to compensate for reduced income from central funds by recruiting more students from the United Kingdom, they may in later years be penalized (*Nature* 22 January). The sources of these contradictory views are easily identified. The marginal cost of extra students is small and many even be zero in some universities and for some courses, so that outside observers may jump to the conclusion that less does not

0028-0836/81/050337-02$01.00

mean fewer. But several universities and colleges are already so hard-pressed that they are being forced to economies that harm their teaching. Laboratory classes are especially at risk. Although the present ratio of staff to students in British universities should not be taken as a magic number, Mr Carlisle's optimism is unjustifiable. And while his statement appears to imply a promise that his own department would be willing to maintain its support for students' maintenance grants at more or less the present level, the capacity of local authorities to keep on paying tuition fees must increasingly be doubted.

The response of the University Grants Committee has been characteristically more cautious. The committee seems more keenly aware of the government's capacity (or incapacity) to pay for the maintenance and tuition of students than is the Secretary of State. Its warning that individual universities should not seek to earn their way out of trouble by recruiting more home students appears to stem from a keener sense of the British government's financial difficulties than from a judicious appraisal of the needs of the university system as a whole. In its circular to universities at the end of last year, the committee said as clearly as is its habit that it will be taking steps, when it makes its financial allocations in the early summer, to control total student numbers and their distribution within universities. One consequence will no doubt be that unseemly competition for students between individual universities will be less open than it might otherwise have been. (But already many universities have partly compensated for the loss of overseas students, driven away by higher fees which were increased by a further 25 per cent last week, by taking more people from around the corner.) If the committee gets its way, there will also be no ostensible change in the present balance between the several universities in the system. No change for the time being seems to be the slogan.

There are two fallacies in this approach. First, it will not work. The competitive universities will see to that. Second, it should not be allowed to work. Further to postpone long overdue changes in the university system will merely make for greater calamity when change is eventually unavoidable. The problem is that the University Grants Committee is constitutionally unable to behave differently. Over the years, the committee has been much admired as a buffer for channelling public funds to autonomous universities in such a way that academic policy is kept out of politics. If a Member of Parliament should raise some question about the purpose of some new sociology course in some smallish university, the appropriate government spokesman will repeat the familiar declaration that academic policy is for universities, that the government habitually accepts the recommendations of the University Grants Committee on how funds should be allocated and that the inquisitive parliamentarian had better write to the vice-chancellor of the university concerned. Even though the committée is not a strictly independent public corporation — it has no bank account of its own, for example — the system worked well in the days of expansion. But the committee, which includes academics representing universities of all kinds, cannot grapple with questions of inequality — should some universities, and if so which, be treated more generously than others? Should some even be denied further support from public funds? Such questions as these are now, however, unavoidable. Who will dare to ask them?

The underlying difficulty, which should command widespread sympathy, is that there is an underlying conflict between the automony of the universities within the British system and the fact that they are supported largely by taxpayers' funds. Both horns of this dilemma are firmly part of British public policy. The independence of the universities is respected and, if threatened, would be fiercely fought for. Public funding, on the other hand, stems from the recognition that there is no other way in which a substantial proportion — for too long too low — of the British population should go on to higher education. The result, however, is that even — or especially — the smallest universities calculate that survival will in the long run be assured only if they ape the biggest and the best. Many of them are ill-placed to follow the successful civic universities, endowed with strong regional loyalties if not with cash, in the development of novel kinds of courses or even local sources of funds.

The most urgent need is that the university system should work out for itself a different pattern at which to aim. All universities are equally free, but that does not imply that they are equal or that they should be dealt with as if they are. Paradoxically, the egalitarian habits of the past few decades have, if anything, served to strengthen the dominance (outside Scotland) of the universities of Oxford and Cambridge. Some other institutions (the London School of Economics and Imperial College, both parts of the University of London, among them) have kept and even strengthened their claims on British intellectual life. Some but not all of the civic universities founded roughly a century ago have branched out in distinctive ways. Some of these occasionally give Oxbridge a run for its money, both in scholarship and the competition for students. Similarly, the group of newer universities includes some with promise, perhaps because of the strength of some department or even because their location helps in the competition for students. The plain truth, however, is that many British universities, for lack of an incentive to behave differently, are undistinguished institutions with unworkable research programmes and with curricula, modelled on Oxbridge, ill-suited to the needs of their students.

The autonomy of the universities notwithstanding, the University Grants Committee is not as powerless to influence this pattern as it seems. Even as things are, a university choosing, for example, to found a medical school would quickly be told that it would have to raise the funds itself, and might even hazard support for its other expenses if it went ahead. Less expensive developments than medical schools are prudently notified to the committee in advance. (Even in these hard times, the committee seems from last month's circular creditably to be planning to earmark some of its budget for new ventures.) There is thus a *de facto* veto on major innovation. The committee is much less able to reshape the pattern of teaching if this entails that some universities should cease doing what they have always done — the committee's attempt to rationalize the teaching of agriculture in the early 1960s, only partly successful, seems to have sapped its courage. Yet present circumstances require that several universities should be asked to stop doing several familiar things.

The simplest change to engineer would be to give up the pretence that all universities are equally able to run graduate schools. This does not, of course, imply that academics need cease to be scholars — as the example of the liberal arts colleges in the United States shows vividly. What it would imply is that small departments of this or that would no longer be given tiny quotas of research studentships by the research councils, would no longer look to the University Grants Committee for the provision of a "research base" — and would no longer face the impossible task of providing a decent graduate education to handfuls of students. The same universities might be compensated by a general understanding, by the committee and others, that they should be places where four-year courses are considered normal, not exceptional. In equity, for the sake of not denying graduates from such universities a chance to make their way in research, the present delegation of the responsibility for choosing graduate students to university departments would have to be stopped. For what it is worth, the committee and the research councils between them have all the power they need to reshape the pattern of British universities in such a way.

How is the committee to summon up the courage for such a task? And how, if at all, is it to choose between the several options in the field? Its own base in the academic world is too narrow, and there is not time for a formal semi-judicial inquiry into the working of the system which it superintends. The first need is that it should decide that a radical change of pattern is needed. Then, given the autonomy of the universities, the committee needs to organize a formal and open (*i.e.* public) consultation among them. The paradox in present circumstances is that, while many academics appreciate the need for change, none of them has an interest that his or her university should make the first sacrifical move in that direction. The committee is the only agency in sight that could free this log-jam.

Johns Hopkins wins new space institute

Princeton loses space telescope ground base

Washington

After racing neck-and-neck to the finishing line, Johns Hopkins University of Baltimore has beaten Princeton University to be the site for a new research institute responsible for the operation of the space telescope to be launched early in 1985. The Space Telescope Science Institute will be managed by the Association of Universities for Research in Astronomy (AURA), a 14-university consortium. The award of the contract to AURA was announced last week by Dr Robert Frosch shortly before he left his position as administrator of the National Aeronautics and Space Administration (NASA).

AURA already operates the group of telescopes on Kitt Peak in Arizona for the federal government. It selected Johns Hopkins as the proposed site last year — even though the university itself is not a member of the consortium and does not at present have a major astronomy department. There were four other bids to operate the institute. Three of these would have put the institute at Princeton and the fourth, from the group which operates the Fermilab near Chicago, would have placed it near the particle accelerator.

The final choice was between AURA and the Princeton proposal from Associated Universities Inc., which already operates the National Radioastronomy Center and the Brookhaven National Laboratory. Dr Frosch accepted the recommendation of an evaluation team that the contract be awarded to AURA. It is thought that a deciding factor was the scope of the management scheme proposed by AURA, based largely on its Kitt Peak experience.

The new institute will be responsible for all aspects of the operation and use of the 2.4-metre optical telescope when it is launched into orbit 350 miles above the Earth from the space shuttle. Johns Hopkins University will provide a site and a new building on the edge of its Homewood campus, at a cost of about $6 million. The AURA plan envisages an eventual staff of 150 people, including 40 full-time astronomers selected from around the world.

In addition, the institute is expected to take in up to 200 visiting scientists each year. They will be able to use the telescope 24 hours a day because, unlike surface-based optical telescopes, the sky will be permanently dark. Annual running costs are expected to be about $10 million. Competition for the management contract has been fierce after bidding was officially opened by NASA at the beginning of last year. There seems, however, to have been little of the political squabbling that has surrounded decisions about major scientific facilities in the past, in particular the controversies over the siting of new particle accelerators.

AURA has made much of the local facilities offered by Baltimore, which has recently been going through something of an urban renaissance. It also stressed the proximity of the city to the Goddard Space Flight Center, from which the manoeuvring of the telescope will be controlled and at which a small team from the institute will be based.

In the end, however, location is thought to have played a relatively minor part in the decision. Princeton is not much further from Goddard, and both cities have access to major international airports — one of the preconditions stipulated by NASA.

More important seems to have been a quantitative assessment made by the evaluation team of the facilities and management support being offered in each case. Johns Hopkins was not able to match the strong astronomy department offered by Princeton. Several of its physicists, however, including some members of the Applied Physics Laboratory (APL), have been closely involved in the design of the space telescope and its instrumentation.

Following NASA's decision, AURA announced that the acting director of the institute will be Professor Arthur Code, at present professor of astronomy at the University of Wisconsin in Madison. The acting deputy director will be Dr Robert Rich of APL. Dr Arthur Davidson, professor of physics at Johns Hopkins, will be acting chief of the research support branch, and Dr William Fastie will be responsible for scientific instruments.

One factor which had concerned the AURA team while the proposals were being evaluated was the threat of federal action against Computer Sciences Corporation (CSC). The consortium had chosen the company to provide computer services, but CSC is under indictment for overcharging the federal government on time-sharing facilities.

Shortly before the management contract was awarded, however, NASA — having checked with the General Services Agency, on whose behalf the indictment has been filed — announced that although CSC has been suspended from bidding for future time-sharing contracts, it was still free to bid for technical support contracts, since these services were provided by another part of the company.

Johns Hopkins officials have welcomed the news about the institute. University president Steven Muller suggested that, given the dominant role that the space telescope is likely to play in optical observation until the end of the century, "there is true excitement at the prospect that Baltimore will now become the world capital of astronomy".

At Princeton, there was comparable disappointment. "To have hosted the institute would have been the next important step for the Princeton scientific community", one faculty member said last week. "Now we must go back and do some reassessing."

David Dickson

Dutch partners for La Palma observatory

The British Science Research Council may have found another collaborator for its optical astronomy observatory at La Palma in the Canary Islands. The Netherlands will probably join Spain, Sweden and Denmark in the collaboration to build and run the site. Last week, the council of the Netherlands Organization for the Advancement of Pure Research (ZWO) approved the deal "subject to contract", as the lawyers say. Professor Graham Smith, director of the Royal Greenwich Observatory which is coordinating the project, welcomes Dutch accession and thinks that the Netherlands may help especially with the supply of skilled manpower for management of the project.

The La Palma observatory, originally called the Northern Hemisphere Observatory, will eventually include our major British telescopes — a 1-m diameter Schmidt telescope, the 2.5-m Isaac Newton telescope (which is being transferred from Britain), a new 4.2-m telescope and a new 15-m radio telescope. Four smaller instruments are being provided by Sweden and Denmark. Preparations for the observatory started in earnest eighteen months ago, after the protracted negotiations with Spain had been completed. Construction is well ahead, according to Professor Smith. The 1-m and the Isaac Newton telescopes are almost ready for removal to the Canary Islands and the observatory is on the verge of placing a contract for the mirrors for the 4.2-m telescope. A call has gone out to astronomers in several countries for the design of ancillary instruments for the 4.2-m telescope.

Successful bidders will be rewarded in part with observing time. The remaining contracts for the 4.2-m telescope are likely to be placed later this year, and the telescope could be working by the spring of 1985, but that date could slip by a year if the cash flow of the Science Research Council does not improve.

Much of the negotiation between the Netherlands and Britain has already been completed. In return for 20 per cent of the cost, ZWO will be allowed to allocate 20 per cent of the observing time to Dutch astronomers. It has not yet been decided precisely how the Netherlands will help to solve the manpower shortage. When the observatory is fully operational, Professor Smith estimates that about 30 technical staff will be needed in La Palma at any one time. The difficulty is in persuading sufficient British scientists to uproot themselves for three years at a time.

ZWO has decided to pay its share out of its existing budget of about £40 million a year. Most of the money will come from the £6 million a year now spent on astronomy, implying a major shift in astronomy funding. Although most Dutch astronomers are expected to welcome the agreement, radioastronomers may feel hard done by.

The next step is for Professor van Lieshout, director of ZWO, to obtain the approval of the Dutch minister for science, who is reported to be enthusiastic. But the imminent Dutch election could mean that the issue will have to be decided by a new minister. All being well, however, Professor van Lieshout hopes that an agreement could be signed and sealed within the next four months.

Judy Redfearn

Toxic waste

Dutch dumps

Amsterdam

The cost to the Netherlands of dealing with chemical waste from years of heavy industrialization is mounting. So far about 3,000 dumps containing chemical waste have been found, 500 of them a recognized danger to public health. The Minister for Public Health and Environmental Protection estimated a few months ago that it would cost about £200 million to clear the 500 dangerous dumps, and this estimate is now £400 million.

In the village of Lekkerkerk, not far from the heavily concentrated chemical industries of Rotterdam, many buildings were found to have been built on a chemical waste dump. About 1,700 drums were recovered from the site, containing materials such as toluene and xylene from the dye industry and metals such as cadmium, zinc and lead. Some 300 houses were evacuated and 150,000 tons of polluted soil have had to be removed, while medical examinations may yet be carried out on the population. Several more chemical waste dumps have since been found in the same area.

After cases of cattle infertility and the discovery of dead birds, an investigation by the municipal environmental laboratory of Amsterdam has revealed that the Volgermeerpolder, a marshy area 5 miles from the centre of Amsterdam, contains about 10,000 drums of chemical waste from a 2,4,5-T factory previously owned by Philips-Duphar. The factory ceased production in 1969 and was completely dismantled and dumped in the Atlantic, but its waste remained.

According to Dr H. Heida, director of the Amsterdam environmental laboratory, the drums contain a wide range of chemicals, including chlorobenzene and chlorophenol. However, some of the drums are known to contain 2,4,5-T and possibly the dioxin 2,3,7,8-tetrachloro-cyclodibenzo-*p*-dioxin (2,3,7,8-TCDD), and it is this latter compound which it is feared may form the real danger.

Analysis of soil, water and livestock and of produce from gardens in the area has revealed the presence of chemicals from the dump in concentrations of tenths to hundredths of milligrammes per kilogramme weight.

However, the future of the dump rests on the problem of clearly identifying 2,3,7,8-TCDD. Samples from the drums have been analysed by Professor O. Hutzinger's team at Amsterdam University and have also been sent to Milan for analysis by the laboratory involved in investigating the Séveso incident. A definite solution to the problem is not expected before the summer.

The clearing of Lekkerkerk

These pollution scandals prompted a national inventory of chemical waste in the Netherlands due to be completed by the end of 1980. So far, however, only two of the eleven Dutch provinces have reported their findings.

The problems of the Netherlands have also attracted international interest. Dr David Costle, administrator of the US Environmental Protection Agency, visited the Lekkerkerk dump while he was in the Netherlands to sign a memorandum on cooperation between the two countries. Dr Tolba of the United Nations Environmental Programme, a regular visitor to the country, is particularly interested in the chemical pollution problems.

The latest revelation was in December of last year, when small amounts of the toxic substance methyl bromide were found in drinking water in the horticultural area between The Hague and Rotterdam. This chemical is used by market gardeners to disinfect the soil, and an estimated 2,000 tons a year are used in this area — 30 per cent of the total consumption in the European Community. The poison had apparently seeped through the PVC tubes used for private water distribution. As a result, the Minister for Public Health has made public water supply companies responsible also for private supply pipes. However, it now turns out that PVC tubes are widely used in the Netherlands for transporting water which has been tested for pollutants such as methyl bromide.

Casper Schuuring

Genetic engineering

Planning bugs

Washington

Echoes of the fierce public debates of four years ago are being heard once again in and around Cambridge, Massachusetts, as local city councils discuss the conditions they will place on the genetic engineering companies springing up in their midst.

The tones of the debate are more muted than before. And in each of the four communities — Cambridge, Waltham, Somerville and Newton — the talk this time is of negotiation rather than confrontation. Nevertheless, there is sufficient concern among the companies for the head of at least one to suggest the need for a "more coherent policy" towards local regulation, possible at the state level.

The most significantly affected so far seems to have been Genetics Institute. This is the company which has been set up by Dr Mark Ptashne, professor of molecular biology at Harvard University, and a local management consultant, Mr Tom Hexner, on privately-raised venture capital, after the Harvard faculty voted against university participation (see *Nature* 27 November 1980). Its backers include Venrock, a venture capital firm owned by the Rockefeller family, and Mr William Paley of Columbia Broadcasting.

Now the company faces a new hurdle — gaining acceptance in the local community. Last November, it applied to build a laboratory in Somerville, just inside the city's border with Cambridge and a few blocks from the university biology laboratories. The application has kindled a fierce public debate. At one point, for example, local citizens had suggested that research should be limited to P1 and P2 physical containment facilities, even stricter than restrictions in Cambridge, which allow work up to the P3 level.

City council members who were originally in favour of granting permission for the new laboratories with few strings attached have backed off in the face of the public controversy; many are preoccupied with severe budget cuts that have resulted from a recent reduction in property taxes.

At a public meeting last Thursday, the

0028-0836/81/050340-02$01.00

council agreed to draw up a local ordinance setting out the conditions under which recombinant DNA research can be carried out in the city. Given the uncertainty, Genetics Institute has now withdrawn its application to build the laboratory.

Mr Hexner said last week that there would now inevitably be a delay in the company's plans. "It really depends on how long the deliberations go on and what the tenor of them is; if the tenor is favourable, then we would probably resubmit", he said. An alternative would be to look for a site in Cambridge itself. In sharp contrast with the debates of four years ago, when the then mayor Alfred Velucchi tried to ban all genetic engineering from the city, an ordinance passed in 1977 is being revised in relatively amicable circumstances.

The revisions have been precipitated by the decision of the Geneva-based Biogen Inc. to construct a new laboratory in Cambridge. Biogen's co-founder, Dr Walter Gilbert, is also a professor of molecular biology at Harvard and was a strong critic of the university's proposed involvement in Genetics Institute.

Biogen is now discussing with the local council the terms under which it will be allowed to operate its laboratory. And the Cambridge Biohazards Committee and the Cambridge Experimental Review Board are discussing a revision to the 1977 ordinance which would require, among other things, that any company or laboratory involved in recombinant DNA research should implement a medical surveillance scheme and a registry of its workers.

Such provisions have already been proposed by the Department of Labor's National Institute of Occupational Safety and Health, on the grounds that any long-term, relatively low-level effects could be traced only if adequate exposure records are kept of laboratory workers and the substances they handle. Biogen has said that it would have little difficulty in meeting such a requirement. More controversial could be the suggestion of a public debate on the acceptability of the site proposed for the laboratory.

Some members of the review board are also hoping to persuade the council to demand independent approval of the safety measures adopted by the company. This might be done by employing a consultant who would report to the city council but whose services would be partly paid for by the company.

A local ordinance including several of the proposed requirements has already been passed by the city council of Waltham, Massachusetts, home of the company Collaborative Genetics. As in Cambridge, the ordinance — which also covers research at Brandeis University — is primarily designed to ensure that all local recombinant DNA research is carried out under the safety guidelines developed by the National Institutes of Health.

In some respects, however, the ordinance is stricter than the guidelines. For example, it requires a formal medical surveillance procedure and the maintenance of work records and also prohibits the use of recombinant DNA products for trials on human patients.

With this model before them, other cities — including Newton, home of New England Nuclear — which are considering their own ordinances may find it difficult to adopt stricter rules. Mr Hexler, for example, says that Genetics Institute would be pleased to live with the Waltham ordinance and that if Somerville is too restrictive, "Waltham is there and waiting".

David Dickson

Lamont director quits

Washington

Columbia University in New York has forced the resignation of the director of the university's Lamont-Doherty Geological Observatory, Dr Mark Talwani, after complaints from senior faculty members about a lack of effective leadership.

Dr Talwani was asked by the university to resign shortly before Christmas. The university's decision was based partly on a report on the situation at the observatory, one of the world's leading centres of oceanic and seismic research, prepared by an outside committee under the chairmanship of Dr Karl K. Turekian of Yale University.

According to the university, the report praised a number of accomplishments of the observatory, but notes the existence of "serious problems" with Dr Talwani's leadership.

The committee itself did not recommend that Dr Talwani be replaced. But after further review and consultation with senior faculty members, including a meeting in October at which "almost all the senior faculty" is said by the university to have supported the search for an alternative director, Dr Talwani was asked to resign within two weeks.

Talwani himself vigorously protested at the university's decision, claiming that it had been the product of a "conspiratorial approach", and complaining to the university trustees that because no specific charges had been levelled against him, "I have had no opportunity to defend myself".

He eventually submitted his resignation as director two weeks ago, although he will remain on the university faculty as a member of its department of geological sciences. A search committee has been established to recommend a successor, and is expected to make its recommendation within a few days. Meanwhile Dr Paul G. Richards, associate director of the observatory, has been temporarily appointed its chief executive officer. **David Dickson**

Soviet fishing

Whither whaling?

The Soviet Far Eastern Fishing Fleet has ceased whaling operations in the Pacific, it was announced in Vladivostok last week. The three factory ships serving the fleet are now being converted into fish-processing plants.

This is the first stage in the implementation of a pledge given at the International Whaling Commission (IWC) meeting in California in December 1978. On that occasion, the Soviet delegate, Vyacheslav Zemskii, pledged his country to phase out all whaling within five years, starting with the North Pacific operation.

The announcement that three factory ships are being converted is, however, somewhat surprising. According to a Moscow radio broadcast shortly before the July 1980 IWC meeting, the Soviet Union had only two factory ships left, the Sovetskaya Rossiya and the Sovetskaya Ukrayina, both sailing from Vladivostok. Both of these ships, the commentator stressed, carried the necessary international observers as well as Soviet marine ecologists.

A partial phasing out of the North Pacific operation began some years ago. In 1979, when the Soviet Union accepted the IWC's partial moratorium, in which fishing from factory ships would be confined to Minke whales, it had already closed down its on-shore whaling stations, and had to negotiate an exchange with the Japanese, ceding its North Pacific quota of 230 Bryde whales in return for 300 of the less valuable Minke whales from the Antarctic.

The closedown of the North Pacific operation does not, however, necessarily mean that the Soviet Union will support the idea of a total moratorium. Soviet commentators have always taken the line that such a policy is not in accordance with the ecological facts, being based on a emotionalism encouraged by some quarters for political reasons. Greenpeace observers have described the Soviet whaling flotillas as near derelict hulks which would cost many millions of rubles to refit. This suggests that the planners have realized for some time that pro-moratorium pressure might well force the Soviet whaling flotillas to disband, and that it was not therefore cost effective to keep the ships in trim.

With the recent drive for self-sufficiency in raw materials, stressed at the October 1980 Central Committee Plenum and in the Guidelines for the new Five Year Plan, what will the Soviet Union use to replace whale-oil?. One possibility is the jojoba, which would probably thrive in the semi-arid steppes of Soviet Asia. Unfortunately, the jojoba plant is dioecious, and, in the natural state, male plants predominate. Jojoba culture is therefore highly labour-intensive, involving the vegetative propagation of large numbers of female

0028-0836/81/050341-02$01.00

plants which then have to be bedded out with a sufficient number of males (the plant is wind pollinated.) With the problem of falling manpower, and the desire to make all agriculture more productive per man-hour, widescale introduction of the jojoba could pose considerable problems of organization.

Vera Rich

Pneumoconiosis unit

Dusty answers

Members of the staff at the British Medical Research Council's Pneumoconiosis Unit in Cardiff are alarmed about the unit's future. So too are the National Union of Mineworkers, the Agricultural Workers' Union and the Association of Scientific, Technical and Managerial Staffs, who say that they value highly the unit's work on industrial diseases. Their fears have been roused by a routine review of the unit's work, conducted by a subcommittee of the council's Physiological Systems and Disorders Board, which has apparently suggested changes at the unit. Last week, council employees and union representatives lobbied a meeting of the council at which the recommendations were to be discussed.

According to the lobbyists, the review has recommended that the unit should continue with most of its present work until a further review in 1987, but that its value should be measured increasingly by the number of contracts for research that it manages to win from industry. They also say that the review has recommended that targets be set for reductions in staff levels, that some work on allergic responses and lung function be relocated and that the unit should not receive a new high-voltage electron microscope, a replacement radiographer nor a replacement respiratory physiologist.

The subcommittee's review is a routine 3-year assessment, to which the work of all council units is subjected. Such reviews frequently lead to decisions to close some units and open others. What seems to have worried the Cardiff staff is that running down some of the work now may pave the way for closing the unit after 1987, when the present director and several other senior staff are due to retire.

Representatives of those lobbying the council's headquarters last week presented their case to the industrial liaison officer who distributed their statements among council members. Their arguments are that relocating some of the unit's work will undermine the multidisciplinary approach needed for many problems and that, by whittling away at the unit's structure now, the council may cause it to become unworkable before formal consultations on steps towards closure can be taken.

The Medical Research Council refused last week to discuss the report of the review committee or the future of the unit. It has still to inform the unit of the decision of the council meeting.

Judy Redfearn

Bulgar birthday

Two for one

"Bulgaria-1300", the unmanned space probe planned as a highlight of the country's 1300th anniversary of statehood, will now consist of two satellites. The first, Bulgaria-1300/1, will be equipped entirely with Bulgarian apparatus, and will carry out ionosphere and magnetosphere investigations. The second, Bulgaria-1300/2, will carry both Bulgarian and Soviet apparatus including several remote sensing experiments, using infrared and VHF wavelengths, and also a special Bulgarian-made multichannel camera. The experiments on Bulgaria-1300/2 will be combined with observations from a high-altitude aircraft and collated with terrestrial measurements.

No details of the experiments have yet been announced. However, according to Dr Kiril Serafimov, chairman of the Bulgarian National Committee for Space Research and Utilization, Bulgaria's "historical circumstances" are such that Bulgarian teams can take part in all five areas of space physics covered by the Comecon joint "Interkosmos" programme — space physics, meteorology, communications, biology and medicine, and remote sensing. New Bulgarian ideas, developed within Interkosmos, stated Dr Serafimov, have included Langmuir probes, ion traps and electrophotometric telescopes. During the past few years, his teams have paid particular attention to noctiluscent clouds and have reported "very interesting data" concerning discharges of auroral type over the magnetic equator. A particular Bulgarian success has been the Spektr-15 multichannel camera, whose 15 channels compare favourably with Landsat's four channels and the East German MKF-6, which has six. (In addition Spektr-15 weighs only 15.2 kg as against the MKF-6 which weighs 158 kg.) The Spektr-15 will be a major feature of Bulgaria-1300/2.

So far, however, there have been no indications of any medical or biological experiments. This seems unfortunate, as the Bulgarian space medicine experiments scheduled for the visit of their Cosmonaut Georgi Ivanov to the Salyut-6 space station were abandoned when his Soyuz transport craft failed to link up with the station. (The other "Bulgarian" parts of the programme, such as the casting of aluminium "foam" using hydrogen as dispersant, were carriked out by cosmonauts Lyaknov and Ryumin, who were already on board Salyut.) Nor are there are any hints that Soviet birthday presents to Bulgaria will respond to the broad hints that the Bulgarian space planners have been dropping ever since Ivanov's mission was aborted — another, and this time more successful, flight of a Bulgarian cosmonaut. Other states have done better, but are not as old.

Vera Rich

Reef aggregation

A new society and a new journal concerned with research on coral reefs have sprung from the frustrations of a group of interested scientists at the difficulties of funding this interdisciplinary work. The new society, called the International Society of Reef Studies, was formed after a meeting held at Churchill College, Cambridge last month.

The first president of the society is Dr D.R. Stoddart from the Department of Geography at the University of Cambridge, who was the leader of the Royal Society's expedition to the Indian Ocean island of Aldabra in the 1960s. The council of the society includes members from mainland Europe, North America, Australasia and the Far East.

One of the objectives of the society — to provide a forum of coral reef research — has been met by the proposed publication by Springer-Verlag of a quarterly journal, *Coral Reefs*, due to appear next year. The new journal will be more concerned with the structure and dynamics of coral reefs than with the flora and fauna, and will thus be complementary to the *Atoll Research Bulletin*, published by the Smithsonian Institution.

The society also plans to take over the management of the international symposia on the subject, of which the fourth is due to be held in Manila in May. Its formation illustrates the difficulties of recruiting funds for the support of research in fields that do not fall squarely within the terms of reference of established grant-making agencies, which have been acutely felt in Britain in recent years. Feeling has run high that governments are interested in coral reefs only in crises.

Brazilian ecology

Polluting by decree

São Paulo

Economic pressures have caused the Brazilian government to postpone legislation banning the use of non-biodegradable detergents which was due to come into force in January 1981.

A decree signed by President Figueiredo last month puts back the starting date for the new law by two years. Industrial pollution is now a serious problem in Brazil. In the state of São Paulo all the rivers are biologically dead and the appearance of foam on rivers is now considered normal, according to the head of the Environmental Technology Agency. Several waste-treatment plants due to open in 1981 in São Paulo will have to be abandoned if the foaming persists.

Although the intended ban on non-biodegradable detergents was drafted four years ago, nothing was done to make it

0028-0836/81/050342-02$01.00

possible to comply with the deadline. Manufacture of the non-biodegradable component of detergents, alquilbenzene sulphonate (ABS), is a monopoly of the Atlantic Richfield subsidiary Empresa Carioca de Detergentes (EMCA). Its production at 45,000 tons a year covers all Brazil's needs. SPUMA, the only company importing biodegradable linear alquilbenzene (LAB) was forced to close down more than a year ago because of prohibitive transport costs from its factory in Manaus in the Amazon region to the main consumption centres of São Paulo and Rio 4,000 miles away. The government chose not to help SPUMA by reducing import taxes and the company's machinery was bought up by a multinational company and left idle.

The Association of Cleaning Products Industries has been accused of lobbying against the entry of biodegradable detergents on to the Brazilian market. The association pleaded for a four-year delay in passing the new law on the grounds that the cost of importing the annual requirement of 35,000 tons of LAB would amount to $35 million in hard currency, increasing the $3,000 million deficit in the Brazilian balance of payments.

What this ignores, however, is the fact that the Brazilian owned company DETEN in now building a plant for the manufacture of LAB in the huge petrochemical complex under construction in Salvador, Bahia. Production is to start in April, reaching 25,000 tons by the end of 1981 and satisfying the expected Brazilian requirement of 70,000 tons per year by 1983. By then the multinational-linked EMCA will also be producing LAB in Brazil at the rate of 25,000 tons per year.

The economic argument for keeping the Brazilian market open to non-biodegradable detergents is further undermined by the statement from the Health Commission of the Brazilian Congress that EMCA had received a firm offer from CONOCO, a Shell subsidiary, to buy up all of EMCA's yearly production of ABS for $36 million for export to other even less ecologically minded countries.

When recommending the President to postpone the ban, the Minister of Industry and Commerce ignored the ecological arguments and cited only the views of the Association of Cleaning Products Industries. The Health Minister in his turn agreed to the postponement on the grounds that health issues were not involved.

The new decree leaves it up to the Ministry of Health to set "indices of biodegradability" over the next two years compatible with the proportions of LAB and ABS on the market. But the Minister of Industry has stated that the responsibility for all future regulations concerning the production, import and distribution of biodegradable products will remain with the executive branch and not with the legislature.

Maurice Bazin

Security conference

Back to Madrid

The Madrid review conference of the Helsinki Final Act reconvened on Tuesday this week for a further two months. Attention now turns from the presentation of reports on progress by the signatory states to the potentially even more acrimonious question of future prospects.

Since the previous review conference in Belgrade in 1977, East–West relations have cooled. One consequence at Madrid was the long procedural wrangle before the conference could begin, partly over whether Afghanistan should be on the agenda.

Later meetings (which were closed to the public) saw much plain speaking, notably from Dr Philip Handler, president of the US National Academy of Sciences. He pointed out that the Hamburg "scientific forum" last February had been devoid of meaning because of Western scientists' concern about "serious infringements of the human rights and freedoms of too many of their colleagues in the East". He concluded with a quotation from Sakharov: "Intellectual freedom is essential to human society — freedom to obtain and distribute information, freedom for open-minded and fearless debate, and freedom from pressure by officialdom and prejudice".

Intellectual freedom concerns only the third of the three "baskets" into which the Helsinki process is conventionally divided. The term "basket" is a relic of the Helsinki Conference in 1973 during which it was decided that there should be three large filing baskets into which delegates could drop their suggestions on, respectively, military matters, trade/technology exchange and human rights.

Western signatories have always maintained, however, that this purely practical procedure did not change the essentially indivisible nature of the process of détente. The socialist bloc, however, would prefer to deal with the three "baskets" in isolation, which, in practice, means concentrating on the military aspects. Indeed, it has often suggested that the Western emphasis on human rights is itself a breach of the Final Act, being an interference in the internal affairs of signatory states.

The discussion of human rights at Madrid has been further complicated by semantics. The United Nations has made a working distinction between "legal" rights and "programme" rights, such as the right to an adequate living standard and the highest attainable standards of physical and mental health. Western signatories insist on the narrower meaning, Soviet spokesmen on the broader.

Proposals tabled for discussion in the next two months are of two kinds, concrete and declaratory. The French, for example, will urge that military "confidence building measures" should be extended to cover the whole of Europe as far as the Urals. At present, notification is required only within a frontier zone of 250 km.

The Soviets, for their part, have called for a European disarmament conference. Poland has proposed Warsaw as a suitable venue, and October 1981 as a possible date. A recent proposal by President Urho Kekkonen of Finland, which is likely to be raised again at Madrid, is known to have strong Soviet backing: the establishment of a nuclear-free zone in Northern Europe. (They are incensed, however, at Danish suggestions that "some regions of the USSR" should be included in the zone.)

Basket Two proposals (trade and technology transfer) will have to be discussed against the background of official trade sanctions imposed on the Soviet Union following its intervention in Afghanistan, and of "personal boycott" pledges by individual scientists and the United States academy's moratorium on exchanges. Major advances are unlikely.

Basket Three will carry a whole range of proposals, from the easing of visa restrictions and the "reunification of families" (the clause of the Final Act most frequently invoked by Soviet Jews wishing to emigrate) to the rights of private citizens to monitor the implementation of the Helsinki accords.

Vera Rich

Indians home from home

Lucknow

In an attempt to recoup some of the talent lost through its relentless "brain-drain", the Indian government has reached an agreement with the United Nations Development Programme (UNDP) to implement a project known as TOKTEN, or the "Transfer of Know-how Through Expatriate Nationals".

Under the project, which it is hoped will further India's National Development Plan, Indian expatriates who are specialists in various technical and sociological fields will be invited to return to India for periods ranging from one week to three months. During their stay contributors would be expected to pass on their skills, act as consultants in their specialist field and make suggestions on issues of policy. They would help put their ideas into practice while in India and continue to give advice afterwards.

Responsibility for running the project will lie with a committee specially formed by the government's Council of Scientific and Industrial Research. Its members include representatives from ministries, government departments, public and private sector organizations, the University Grants Committee and the Indian branch of UNDP. Meanwhile UNDP has allocated an initial $100,000 for travel and living expenses.

So far about 300 Indians — including doctors, engineers, scientists and social scientists — have agreed to take part in the TOKTEN project.

Zaka Imam

0028-0836/81/050343-01$01.00

CORRESPONDENCE

Galileo's trial

Sir — Your recent leading article "Why bother to rehabilitate Galileo?" (*Nature* 30 October 1980; **287**, 767–768) begins with a pertinent review of Galileo's trial. The last part of the article unfortunately becomes an arbitrary critique of the twentieth century Catholic Church which does not belong in a scientific journal like *Nature.*

More important though, some errors are present. For instance, it is stated that directed evolution plays a "big part in the Church's view of biology", that the Big Bang is "leaped at" by the Church and that the scientific failure of these two theories would place the Church in an indefensible position. These three statements are invalid.

The official Church never taught "directed evolution". Teilhard de Chardin never was a "doctor of the Church" and the official theology of the Catholic Church is still founded on the works of men such as St Thomas Aquinas, who admitted the possibility that the world apparently acts as moved by the laws of fate and necessity (see also St Augustine).

The Big Bang was once kindly considered (but never affirmed) by Pope Pius XII. And some modern astronomers actually claim that the Big Bang is a proof of God (which is either stupid or blasphemous, and perhaps both). The finiteness or infiniteness of space, time and the number of intelligent beings in the Universe are not fundamental problems and have no bearing on the primary question "Why existence?". Science and God, for the Catholic Church, do leave the individual free to answer such questions for himself, and it seems useful that you should be aware of the affirmation of Pius XII, Paul VI and John-Paul II that science and faith each have their own field.

The article "Why bother to rehabilitate Galileo?" ends with categorical criticism of the Church's moral (sexual) recommendations. What is the point of this? The Church proposes a way of life (responsible paternity) to people who believe in the Church and in the eternal destiny of men. Ethics is not a consequence of science.

Then you continue to say peremptorily that "The Church's views on birth control . . . . are a means of making millions of small-time martyrs, pathetic contemporary analogues of Galileo" — what can be said of abortion, which involves an act similar to that of the men who condemned Galileo?

Henri Reboul

Université des Sciences et
Techniques du Languedoc,
Montpellier, France

Don't be vague

Sir — Mr Ronald Reagan may indeed be conservative but he is surely not so backward-looking as to choose a British First World War general as his Secretary of State (*Nature* 18/25 December 1980, p.631).

David Pike

King's College Hospital,
London SE5, UK

Secretary of State Alexander Haig was inexcusably confused with Earl (Douglas) Haig (1861–1928) — Ed.*Nature*

Evolutionary thought

Sir—At a time when irrationalism is gaining ground against science in popular culture I feel that I must echo some of Dr Halstead's misgivings about the use of cladistics to explain the process of evolution[1]. However, his castigation of Engels as the *eminence grise* behind cladistics is unfounded and is based upon a misunderstanding of how dialectical materialism explains discontinuous development in both nature and society.

Halstead (with Popper) sees Marxism as "working to spark off the revolution"[2]. This is the sort of ultra-leftism espoused by the early German Communist Party and later by Stalin in his "Third Period". Lenin in "Left-wing Communism — an infantile disorder" answered the former and Trotsky the latter in "Revolution Betrayed". In the biological sciences, organic change may be either gradual or discontinuous. Darwin by over-reacting to "leaps" in development, which he saw as the "new creations" of his religious opponents, ended up by formulating a theory of inheritance based on the blending of acquired characters — pangenesis. This ultra-gradualism was overthrown by genetics which reinstated the view of saltatory leaps, now based on mutations. Dialectical materialism properly understood proposes neither an opening for deism nor mutational leaps. It does propose the transformation of quantitative developments (genetic mutations and changes in ecology) into a qualitative leap in development (not in structure but in selection pressure). A genetic mutation *per se* is a leap, but not a dialectical leap.

The clearest examples of dialectical change in evolution are in the area of so-called macroevolution, where quantitative genetic changes taking place to fit an organism for one habitat pre-adapt the organism to "leap" into another ecological niche where the selection pressure is qualitatively different. Darwin describes this type of development in the speciation of the Galápagos finches, although not referring to it as a leap. The early diversification of the mammals has been described by Olson[3] as due to "Co-ordination and integration of such modifications (as) brought some members of a particular radiation to the threshold of a new radiation". Similarly, many workers explain the initiation of major periods of diversification or the invasion of new environments by such a "quantitative to qualitative and back to quantitative" hypothesis.

Even G. G. Simpson, to whom the mechanistic view of cladism may or may not have been anathema, was by no means averse to the idea of leaps. I quote: "Even within the staid horse family, which seems as a whole to be progressing rather steadily through the Tertiary, close examination shows the rates varying considerably. More broadly, evolution commonly seems to proceed in spurts and pauses in an apparently erratic way"[4].

Mike Howgate

Department of Zoology,
University College London,
London WC1, UK

1. Halstead L. B. *Nature* **288**, 208 (1980).
2. Halstead L. B. *New Scient.* 17 July 1980, p.215.
3. Olson E. C. *Vertebrate Paleozool*, 673 (Wiley-Interscience, 1971).
4. Simpson G. G. *The Meaning of Evolution*, 103 (Yale University Press, 1976).

Engels on Darwin

Sir — How right Dr Halstead is to remind us of the lengths to which Marxists have gone, and will no doubt continue to go, in their efforts to place their materialist theory of history on a scientific par with Darwin's evolutionary theory. Engels, no less, argued, in his oration at Marx's funeral, that ". . . just as Darwin discovered the law of development of organic nature, so Marx discovered the law of development of human history. . . "

Incidentally, it is worth recalling that the only Englishman at the graveside to hear this, was Marx's erstwhile young friend Ray Lankester FRS. Lankester was later knighted and served as the Director of the Natural History Departments of the British Museum (1898–1907), and is still remembered for his reclassification of the Talpidae collection. The plot thickens!

Harry Crabtree

Birmingham 13, UK

Motorbike safety

Sir — A recent letter to *Nature*[1] concerns the observation that an object appears to maintain constant size as its distance from the observer changes. It is simply "seen" as a full-sized object at varying distance. This size constancy fails quite suddenly when the object subtends an angle at the observer of less than half a degree; it is then seen as a small object, presumably of less immediate importance.

This phenomenon could be a contributory factor in a fairly common type of road traffic accident, that caused when a driver pulls out unexpectedly in front of motorcyclist who is then unable to stop.

An approaching motorcyclist presents an irregular silhouette subtending an average angle of less than half a degree at any distance greater than about 200 feet. If size constancy fails below half a degree, the motorcyclist would only be seen as small, and would therefore be less likely to register subconsciously as a real full-sized danger, until within about 200 feet.

The shortest possible stopping distances for a moving vehicle under good conditions are reported in the highway code as follows

Vehicle speed (m.p.h.)	30	40	50	60	70
Stopping distance (ft)	75	120	175	240	315

Interpolation suggests that when a motorcycle is approaching at over 54 m.p.h., so that its stopping distance is greater than the critical 200 feet, an accident might be inevitable, particularly if either driver was not concentrating on "seeing" the other vehicle at the crucial time. Many similar calculations are possible. A convex rear view mirror reduces the image size and so can make it possible for larger vehicles to approach within their stopping distances before they, too, subtend the critical angle of half a degree.

Much work has been done on the effect of factors such as brightness, contrast, colour, shape, size (and styling features disguising shape and size) on the visibility of an object. The results of this work might valuably be applied to problems of the type outlined above

G. K. McGinty

Redhill, Surrey, UK

1. Ross J., Jenkins B. & Johnstone J. R. *Nature* **283**, 473–474 (1980).

The Nobel awards for medicine or physiology

from Sir Peter Medawar CH, FRS

THE joint award of the 1980 Nobel prize for medicine or physiology to Jean Dausset, George Snell and Baruj Benacerraf has been well received by immunologists and medical biologists generally. The general public and the scientific press also found the award easy to understand: Jean Dausset was prime mover in the discovery in human beings of an important system — 'HLA' — of tissue groups that affect the outcome of organ transplantation in a way analogous to that in which the principal blood groups — ABO — affect the outcome of blood transfusions.

Its relevance to tissue transplantation is not, however, by any means the most important aspect of the discovery of the so-called 'major histocompatibility complex' (MHC) of human beings and of the analogous — perhaps homologous — MHC of mice, known as H–2.

The great and abiding importance of the HLA system is twofold. First, it defines a system of polymorphism — of stable genetic differentiation — in human beings which has uncovered, as no other procedure could have done, the genetic basis of differences in susceptibility to multiple sclerosis, ankylosing spondylitis and the juvenile (insulin-dependent) form of diabetes. Even more important, it has become increasingly apparent that the MHC of man as of the mouse defines a whole system of cell-surface markers which are of the utmost importance for the organism's power to discrimate between self and non-self; indeed, an organism's power to react to a number of immunogens as antigens seems to depend on their being signalled by the gene products of the MHC.

Although Jean Dausset was prime mover in the recognition and definition of the HLA complex, this discovery — as with all others of comparable stature — was the work of a number of medical scientists standing shoulder to shoulder: the names that come instantly to mind are those of Bernard Amos, Richard Batchelor, Walter Bodmer, Ruggero Cepellini, Rose Payne, Jan van Rood and Paul Terasaki — a splendid example of international co-operation promoted by the Transplantation Society's generosity and good sense in organizing histocompatibility workshops at which all the principals met to exchange ideas, knowhow and reagents.

To explain George Snell's award we must go back to the early days of tumour transplantation. In the first decade of this century it was widely believed that a study of the factors promoting or retarding the growth of transplanted tumours held promise of a prevention or cure of malignant growth; but by the end of the decade, thanks mainly to the criticism of Peyton Rous, it came to be recognized that the study of tumour transplantation was a study of transplantation rather than of tumours, for in the days before homozygous strains of mice became available, tumour transplants were all of the kind known as 'allografts' and as such aroused immunological rejection reactions.

A very great deal of the work was done with what might be called 'any old mice'; it made no kind of sense and — what is worse — still cannot be interpreted today. This judgement applies with especial force to work on 'immunization' against transplanted tumours by the previous inoculation into their intended recipients of tumour cells of the same kind or of a variety of normal tissues.

Matters were in a desperate state until they were taken in hand by Little, Strong, Bittner, Johnson, Bagg and George Snell of the Roscoe B. Jackson Memorial Laboratory — the men who originated the inbred strains used throughout the world today. Using genetically uniform mice it became possible to formulate the elementary genetic ground rules of transplantation of normal tissues or of tumours. George Snell went further and worked out in great detail how the outcome of transplantations is governed by the genes that comprise the MHC of mice (H–2), first serologically recognized by Peter Gorer of Guy's Hospital, London, who if he had lived, would certainly have shared the Nobel prize. The study of H–2 provided the technology and the conceptual background that made possible the recognition of the MHC of man.

This has been the international year of the MHC, for the award to Baruj Benacerraf also relates to the MHC — in particular with that sub-region of it that has to do with '*Ir*' (short for immune response) genes. When Benacerraf and H.O. McDevitt first recognized this sub-region it was widely thought that they had discovered the gene that coded for the antigen-specific T-cell receptor, but it now seems more likely that the *Ir* region is involved with the presentation of antigen to T-helper cells.

The work of this year's recipients of the Nobel award for medicine or physiology has a twofold lesson to teach. The first is of the immense fruitfulness of collaboration between mice and men in the solution of medical biological problems. The second, not quite so obvious, is that the discovery of HLA provides an exemplary case history in rebuttal of the notion that scientific discovery can be premeditated. What advice, for example, should one give to a young man intent on finding out the genetic basis of differences in susceptibility to multiple sclerosis, ankylosing spondylitis or insulin-dependent diabetes? It would be the *reductio ad absurdum* of the notion that scientific discovery can be premeditated to recommend that the young research worker should embark on a study of tumour transplantation, specifying the use of heterozygous mice, until its contradictions and vicissitudes drove him nearly mad whereupon he should proceed at once to work out its genetic and serological basis, from which the results could be translated to human beings and the rest would follow.

In days when businessmen and others who ought to know better are advocating a diminution or even abandonment of 'pure' research in favour of its practical applications it cannot be too strongly emphasized that discoveries such as HLA — as of diagnostic X-ray radiography — were made by the ordinary processes of scientific discovery, slow, messy and uneconomic though they undoubtedly are. It is in fact only unworldly impractical daydreamers who think that scientific discoveries can be premeditated or can be arrived at by the working of some calculus, some organon of discovery. It was the tough practically minded, no-nonsense men of affairs — particularly of great affairs — the world's Rockefellers, Sloans, Nuffields and Wolfsons who proved to understand most clearly the need and to provide most generously the means for what is so mistakenly called 'pure' research. □

Sir Peter Medawar, CH, FRS is Head of the Transplantation Biology Section of the Medical Research Council, UK.

0028-0836/81/050345-01$01.00

Meteorite breakup

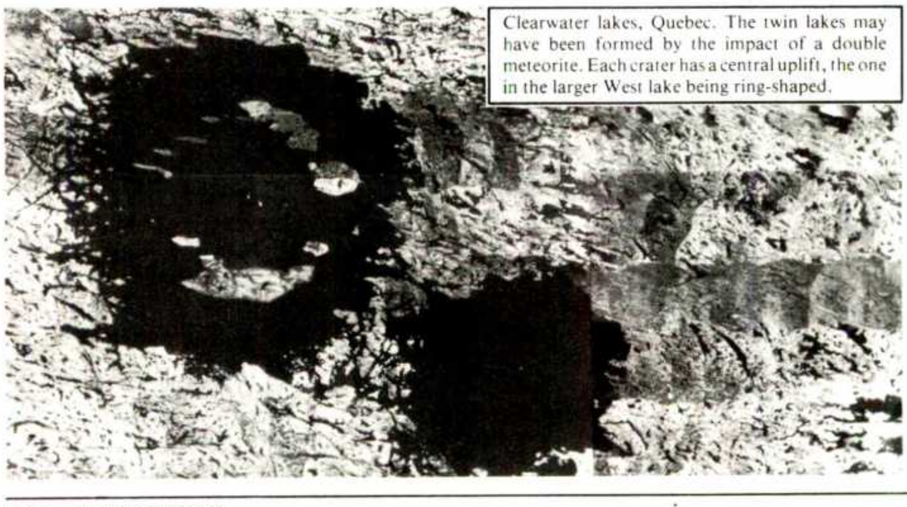
from Michael R. Dence

INTEREST has revived in crater groupings on the solid planets since the report that some asteroids may be double (Tedesco *Science* **203**, 905; 1979). Earlier debates between those who considered that groups resulted from two or more simultaneous impacts and those who considered them to be chance clusters of unconnected events have run in favour of chance (see Woronow *Icarus* **34**, 324; 1978). Yet some crater associations, most notably on Earth, seem to require coeval impact events. Can all such groups be accounted for by breakup of single bodies before impact or are they evidence for meteoroid pairs in space? To answer this question statistical studies have now been augmented by attempts to derive a quantitative understanding of meteorite disruption by an atmosphere or by tidal rupture.

The breakup of meteoroids or of re-entering satellites has been described by many eyewitnesses. Disintegration of meteorites has also been photographed and led in the case of the Innisfree Alberta fall of February 1977 to the recovery of nine fragments of the aerolite (Halliday, Blackwell & Griffin *J. R. astr. Soc. Can.* **72**, 15; 1978). Much larger meteorite showers may contain many hundreds of fragments. In such falls the meteorite fragments, or the craters they produce if they land with sufficient velocity, are usually distributed over an area elongated in the direction of flight of the bolide. The strewn field or scatter ellipse is commonly only a few square kilometres in area, but, as in the case of the great 8 March 1976 fall at Kirin, North China, may cover several hundred square kilometres. Although, on average, the scatter ellipse is one-third as wide as it is long, some are almost circular while others show a down-range dispersion more than eight times the cross-range scatter. Commonly, but by no means invariably, the largest fragments or craters are found at the down-range end of the ellipse.

What causes this scatter? The general principles have been understood for some time and a few falls have been analysed in considerable detail. The veteran Russian meteoriticist E. L. Krinov concluded that the massive Sikhote-Alin fall of 12 February 1947 came from an iron meteorite weighing more than 70 tonnes which broke up in three stages. From the pattern of secondary scatter ellipses superimposed on the main ellipse, Krinov deduced that there was an initial high-altitude breakup followed by two additional disruptions, the last at an altitude of 6 km or less (*Meteoritics* **9**, 255; 1974).

This and other large falls, mostly of iron meteorites, have now been modelled quantitatively by Passey and Melosh (*Icarus* **42**, 211; 1980) with some unexpected results. They trace the history of individual meteoroids of varying size and crushing strength after they enter the atmosphere at an altitude of 70 km. Velocity and entry angle are also varied in the calculations. Initial fragmentation is taken to be largely the result of aerodynamic pressures on pre-existing defects or planes of weakness, though they recognize that thermomechanical stresses may also be effective. The mechanics of breakup are not fully understood and the estimation of appropriate properties such as crushing strengths is one of the main uncertainties of the analysis. Following breakup, dispersion down-range is considered to result mainly from atmospheric drag. Since the small fragments are decelerated more effectively than the large, they tend to fall short of the main mass. Lift may counter the effect of drag, but Passey and Melosh argue that it is only effective for fragments lighter than 10^2 kg. For larger masses they conclude that lift is a small fraction ($\sim 10^{-3}$) of drag, otherwise many more would lie down-range of the main mass, or crater, than is generally observed. Parenthetically it may be noted that Sikhote-Alin is one instance where many sizeable craters are down-range from the largest.



Clearwater lakes, Quebec. The twin lakes may have been formed by the impact of a double meteorite. Each crater has a central uplift, the one in the larger West lake being ring-shaped.

A feature of their analysis is the attention Passey and Melosh give to cross-range dispersion. This they attribute to a combination of the effects of bow shock interaction, spin of the meteoroids acquired before entry, and spreading induced by crushing as the meteoroid decelerates. All of these effects they judge to be quantitatively similar and capable of explaining most of the observed scatter transverse to the direction of flight. They calculate that maximum separation is achieved when breakup occurs at an altitude of about 15 km.

More of a problem is their conclusion that, for meteoroid masses of 10^4 to 10^9 kg, multiple crater fields result when the angle of entry of a bolide is less than 30°. Indeed, in analysing specific terrestrial crater fields, formed for the most part by iron meteorites, they generally favour entry angles of less than 20°. However, the examples they draw on include 70 per cent of the terrestrial inventory of craters 20 to 200 m across with associated meteorite fragments. As the probability of the entry angle being 30° is only one in four and is even less for lower angles, it is evident that the result implies a marked deficiency of craters from high-angle impacts. As there are no other grounds for believing this to be the case, their analysis requires further examination. One method of expanding the set of conditions for producing crater fields is to lower the strength assumed for the meteoroids. Passey and Melosh adopt yield strengths between 10^7 and 5×10^8 Nm^{-2} in their representative calculations. This range may be appropriate for the low-altitude, final breakup of the Sikhote-Alin meteoroid, but its initial, high-altitude disruption implies considerably lower strength for the main mass. We are brought again to the need for an improved theory of meteorite fragmentation.

Further problems arise when bodies of mass greater than 10^9 kg are considered. Basically, much larger bodies on entering the atmosphere may rupture but no combination of atmospheric deceleration, lift, bow shock interactions or gravity will allow sufficient separation for the formation of more than one crater. For typical velocities of 15 to 20 km s^{-1}, separations of the order of 20 times the radius of the impacting bodies are required to achieve clear separation of the resulting craters. Nonetheless, pairs of large craters, formed by bodies of the order of 10^{12} kg, are found on Earth. The best known are the 32 km and 24 km diameter pair at Lac à l'Eau Claire (Clearwater Lakes, see figure), northern Quebec, Canada and the 24 km Rieskessel and 3.5 km Steinheim Basin combination in southern Germany. The former have a centre-to-centre separation of 31 km, the latter of 45 km. Radiometric

Michael R. Dence is Acting Director of the Gravity and Geodynamics Division of Energy, Mines and Resources, Ottawa.

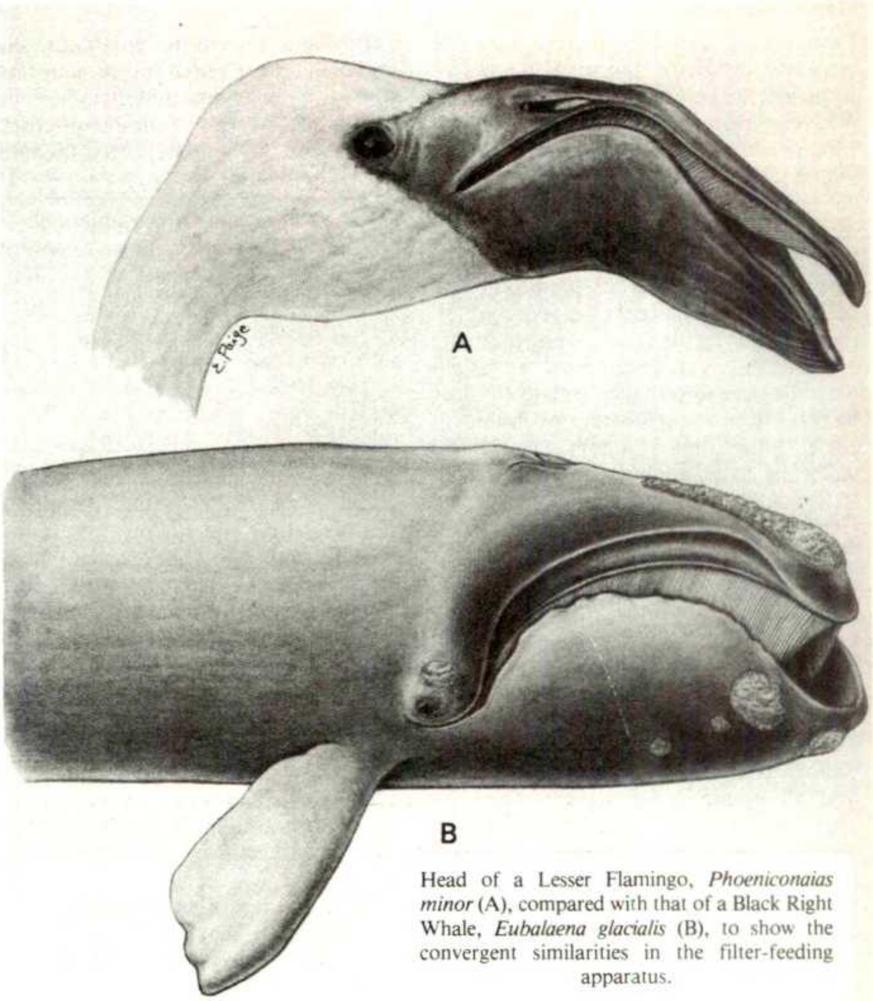
0028-0836/81/050346-02$01.00

Flamingos, stilts and whales

from Andrew Milner

THE phylogenetic position of the flamingos (family Phoenicopteridae) has long been the subject of controversy among avian systematists with some workers supporting relationship with the Anseriformes (ducks and geese) whilst others favour association with the Ciconiiformes (storks, herons and ibis).

The evidence has been reviewed recently by Olson and Feduccia (*Relationships and Evolution of Flamingos (Aves: Phoenicopteridae)* Smithsonian Contributions to Zoology **316**, Smithsonian Institution Press; 1980) who have developed an earlier proposal by Feduccia that the flamingos fall within the Charadriiformes (waders, gulls and auks). Characteristics of the musculature, the skeleton, the natal down and the endoparasitic cestodes all lead to the conclusion that the closest living relatives of the flamingos are the Recurvirostridae (stilts and avocets) and, more specifically, the Australian Banded Stilt (*Cladorhynchus*). This remarkable bird lives in colonies frequenting temporary salt lakes in southern Australia and its behaviour, pattern of breeding and life history all bear specific similarities to those of flamingos. The earliest certain fossil flamingo from the Eocene of Wyoming is described by Olson and Feduccia and supports their hypothesis by being morphologically intermediate between recurvirostrids and flamingos. One phylogenetic anomaly, the presence of anseriform-type feather lice in flamingos but not in *Cladorhynchus*, is attributed to the presumed common lifestyle and physical proximity of early flamingos and the early colonial Anseriformes such as *Presbyornis*.



Head of a Lesser Flamingo, *Phoeniconaias minor* (A), compared with that of a Black Right Whale, *Eubalaena glacialis* (B), to show the convergent similarities in the filter-feeding apparatus.

The filter-feeding mechanism of flamingos is structurally and mechanically different from that of ducks but is remarkably convergent with that of the baleen whales, particularly the Right Whale! Despite the differences in scale and orientation during feeding, the two filter feeders share structures such as a large fleshy tongue accommodated in a deep-sided lower jaw and a recurved rostrum permitting the presence of a row of elongate filtering devices suspended from each upper jaw. The similarity in structure and function of the feeding apparatus of whale and wader suggests to the authors that the 'crooked beak' of the flamingo may have been a fundamental adaptation to filter-feeding which preceded the 'inverted-head' feeding posture rather than following it as has generally been assumed.

Andrew R. Milner is in the Department of Zoology, Birkbeck College, University of London.

---

and geological dating leaves little doubt that each pair was formed simultaneously. In addition, in neither is there a structural basis for assuming a highly oblique impact angle. It seems that both pairs may indeed be the product of the impact of double meteorites.

The idea is in fact not new having been examined by R. W. Tanner in 1963 (*J. R. astr. Soc. Can.* **57**, 109). Recently it has been considered in greater detail by P. D. Noerdlinger (*Icarus* in the press) who, after an exhaustive study of tidal disruption mechanics, confirms that Tanner's analysis is essentially correct. In the general case of tidal rupture at a distance of several Earth radii the separation that fragments achieve approximates only the initial radius of the parent body. Overlapping or nested craters will result. It is conceivable that the separation found at the terrestrial crater pairs might be achieved by tidal or atmospheric rupture of a body of low strength which, entering at an oblique angle, became a satellite for several orbits. Clearly such an event has an extremely low probability, but the Ries-Steinheim and Clearwater pairs are two occurrences in a total population of approximately two dozen large ($\sim$ 20 km) craters (Grieve & Robertson *Icarus* **38**, 212; 1979). Once again the existing crater statistics argue that multiple cratering events, even in this size range, are not of great rarity.

On the other hand Noerdlinger contends that the population of meteoroid doubles is also probably small. Rupturing an asteroid so that the pieces remain in close association requires a delicate balance between the energy of fragmentation, gravity and angular momentum. The association may also be difficult to maintain for any appreciable time. The entire life cycle of double meteoroids and asteroids is in need of a careful, quantitative evaluation of the circumstance under which they can form and their subsequent stability before final capture, separation or disintegration. Such a study would be relevant to another important issue — the formation and lifetime of ring systems as now revealed in spectacular detail by the Voyager images of Jupiter and Saturn.

0028-0836/81/050347-01$01.00

Cement — a respectable material?

from D.D. Double

FAMILIARITY breeds contempt. We see cement everywhere, mixed with sand to form mortar between bricks, mixed with pebble aggregate to make concrete. World-wide production of cement approaches a billion (10^9) tonnes per annum and of this 95 per cent is Portland cement, the remainder being more specialized products such as high alumina cement. This figure very substantially exceeds the production of steel, traditionally its nearest counterpart in large-scale engineering construction. Yet, in spite of its obvious technological importance, relatively little has been done to explore the possibilities of cement other than as a convenient gluing agent in a cheap bulk filler material. Relegated to this rather unsophisticated role, outside the established industry, cement has not received the attention that it deserves.

But this situation may be changing. In general, there is increasing interest in inorganic materials that can be produced cheaply and which might compete with or even replace other common materials which are more costly in economic or ecological terms[1,2]. In this context, silicate- and aluminate-based systems seem an attractive proposition and one of the obvious sources for development is cement (principally Portland but also other hydraulic cements). Materials scientists, who have hitherto mostly ignored cement as a not-quite-respectable material, are now taking a different view.

Portland cement is remarkably cheap and materials required for its manufacture (limestone or chalk, and clay) are widely abundant. In terms of the energy costs for production it compares very favourably with other materials.

Energy cost for production
(per unit volume, relative to cost of cement)

Portland cement	1
Glass	1.3
Aluminium	35
Steel	20
Polystyrene	5

Whereas other materials, such as glasses, ceramics and metals, usually require high-temperature, and therefore expensive, shape-forming processes, cement develops its strength by reaction with water at ambient temperatures. Before setting, it is easily castable and made into complex shapes. With a density around 2.4, compared for example with that of aluminium 2.7, a hardened cement paste is relatively light. It has low thermal and electrical conductivity, is inherently non-combustible and, as far as one can tell, is generally non-toxic. Young's modulus has quite a respectable value around 20 GN m^{-2}.

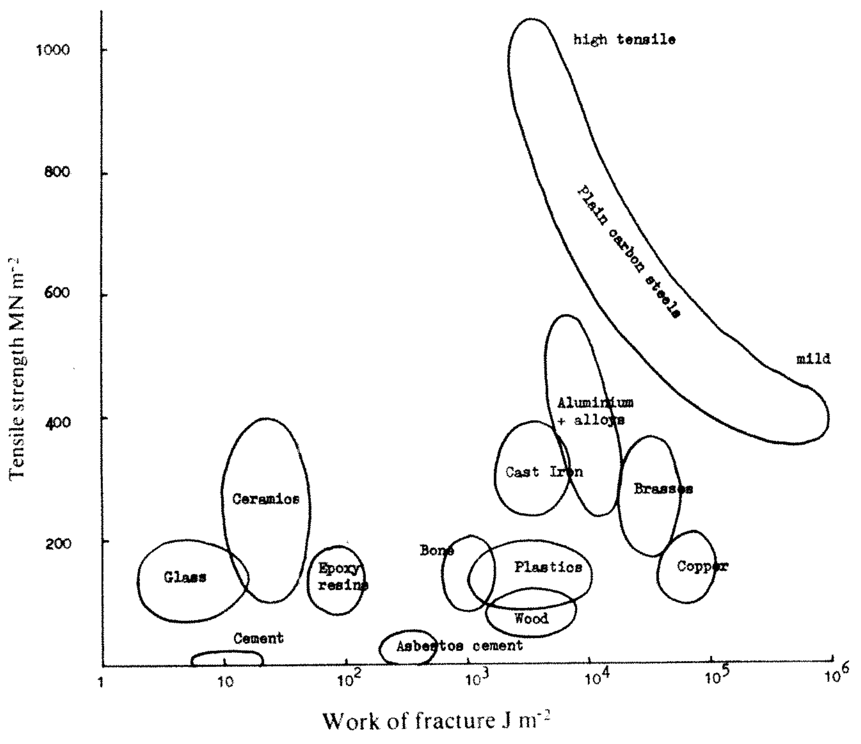
With all these potential advantages, one might wonder why cement has not been used more widely in more diverse ways. The reason is quite simple — as normally made, it is very weak and brittle. Although a hardened cement can sustain quite appreciable compressive loads, its tensile strength and fracture toughness — two parameters that often determine the practical use of a material — are disappointingly low, respectively 5–10 MN m^{-2} and about 20 J m^{-2}. The situation in relation to other common materials is illustrated in figure 1. Engineers are well aware of the mechanical deficiencies of cement as a binder. Wide safety margins are usually used when dealing with structural concrete sections and methods are devised (reinforcing, prestressing, cantilevering) to avoid putting concrete under any substantial tensile stress.

It is obvious that if cement is to be exploited in new and unconventional ways then improvements have to be made in the mechanical properties. Cement will never be as strong as steel but, referring to figure 1, one can see that only relatively modest improvements are required to bring it on a par with a range of other materials. Cement could become a serious contender at a tensile strength of 100 MN m^{-2} and 10^3–10^4 J m^{-2} in a work of fracture. The question is whether such improvements are possible.

How does nature deal with the problem of forming strong inorganic solids in an aqueous medium at normal temperatures? The classic example is the calcium carbonate found in shells of marine organisms. Measurements of mother-of-pearl shell (nacre) show that it has a modulus around 100 MN m^{-2}, a tensile strength between 70 and 100 MN m^{-2} and a fracture energy greater than 1,500 J m^{-2} (ref. 3). It is clearly a much more impressive material than $CaCO_3$ in the form of chalk, limestone or even the best quality marble. It achieves these properties by developing an intricate structure of low porosity consisting of plates of crystalline $CaCO_3$ (aragonite) held together by very thin layers of a protein polymer 'glue' (figure 2). Limpet shells have a more complex cross-laminated structure and certain crab shells show an interesting, and presumably purposeful alternation between calcitic and aragonitic crystal forms of $CaCO_3$. These examples demonstrate that, through appropriate reorganization of the micro-structure, a rather unpromising substance can be transformed into a mechanically sound material. The same principle should apply in the case of cement.

In general, hydraulic cements harden by reaction with water through precipitation of near-insoluble hydration products which progressively fill the spaces between the residual cement grains, forming a cohesive matrix that binds the composite mass together. In Portland cement, the main binding agent is a colloidal calcium silicate hydrate gel of rather indefinite

Fig. 1 Tabulation of tensile strength versus work of fracture for a range of common materials.



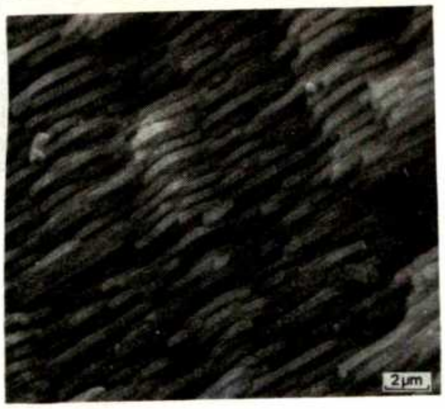


Fig. 2 Fracture surface of mother-of-pearl shell (nacre) showing intricate laminated structure of calcium carbonate (aragonite) crystals. Scanning electron micrograph. (Photograph courtesy of J.P. Northover, Oxford University.)

composition and structure. In fractured samples this gel is often seen as irregular interlocking fibres growing from the surfaces of the cement grains and interspersed with plates of crystalline calcium hydroxide, a by-product of the hydration reaction (figure 3). The mechanism by which hydration occurs and the factors responsible for the development of strength are still not fully understood (see for example ref. 4). It is recognized, however, that the main source of weakness is the residual porosity within the hydrated cement microstructure. As normally made, a hardened cement paste has a total porosity of about 20 per cent and contains pores with sizes ranging from the order of 1 mm to less than 100Å.

The importance of the article in this issue of *Nature* (p.388) by Birchall, Howard and Kendall is the convincing demonstration that, by deliberate manipulation of the porosity, very substantial improvements can be made in the strength of cement. The flexural strength values quoted, between 60 and 70 MN m^{-2}, are about an order of magnitude greater than those for 'normal' cements, and even exceed previous estimates of the maximum intrinsic strength.

This has been done by application of classical theory relating to the fracture of brittle solids (Griffith's theory) and by recognition of the fact that it is porosity distribution that determines the fracture strength and not total porosity, as has hitherto been supposed. Although the total porosity of their samples remains fairly high, the increases in strength have been achieved by refinement of the porosity distribution to eliminate the large pore range — the macrosopic flaws that provide sources for easy propagation of cracks through the microstructure. Thus, the article shows that cement is not intrinsically as weak as has been supposed and also illustrates that the improvements are attainable by fairly simple methods without the need to apply elaborate and costly treatments such as hot pressing, autoclaving and resin impregnation — methods that have been investigated by previous workers.

This may be only the beginning of the story. After physical processing to reduce the porosity to dimensions comparable with the size of the cement grains and the hydrate particles, researchers should perhaps be looking to methods of chemically modifying the hydration reaction in such a way as to promote greater cohesion within the cement paste. Previous attempts to do this have not been successful probably because of the overriding effect of macro-porosity. Chemical cross-linking by the incorporation of reactive organic molecules is one possibility; an enhancement of the natural polymerizing ability of the colloidal silicate in Portland cement is another. The chemistry is extremely complex but, nevertheless, there would seem to be considerable scope in this area. Fibre reinforcement is a well proved method of augmenting the mechanical properties of a material, particularly the fracture toughness. It has been applied with some success in the case of cement using asbestos, glass, steel and polymer fibres, but in many cases the mechanical deficiences of the cement matrix have imposed limitations. An enhancement of the inherent strength of cement could provide considerable impetus to the development of such composite materials.

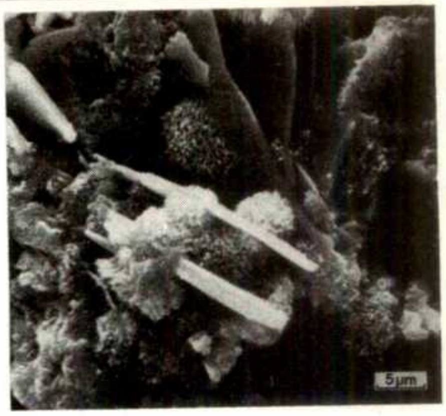


Fig. 3 Fracture surface of hardened Portland cement paste, showing fibrillar growth of calcium silicate hydrate gel from surfaces of cement grains and interspersed crystals of calcium hydroxide. Scanning electron micrograph.

There is still a long way to go but if the necessary improvements are possible, it is difficult to predict where cement-based materials might eventually be used. One can visualize a variety of new domestic uses and even in the area of transport. Boats have been made out of reinforced cement (ferrocement) for many years. Cement, or something rather like it, may yet fly — not as a turbine blade in a jet engine certainly, but perhaps as an important structural part of the airframe. If that day comes, then cement will have become a very respectable material indeed. □

D.D. Double is in the Department of Metallurgy and Science of Materials, University of Oxford.

---

1. Birchall, J.D. *Proc. Symp. Adv. in Cement Matrix Composites* (Materials Research Society, Boston, Massachusetts, 1980).
2. Ray, N.H. *Inorganic Polymers* 150 (Academic, New York, 1978).
3. Currey, J.D. *Proc. R. Soc.* **B196**, 443 (1977).
4. Lea, F.M. *Chemistry of Cement and Concrete* 3rd edn, Ch.9 & 10 (Edward Arnold, London, 1970).

Understanding interatomic forces

from Gillian Peach

THE ANALYSIS of spectra often provides the only diagnostic tool with which to study the physical conditions of a plasma or gas in the laboratory or in space. The interactions between the atoms that are emitting (or absorbing) radiation and neighbouring atoms can, however, produce a marked effect upon the observed spectrum. In 1907 R.W. Wood (*Astrophys. J.* **26**, 41) wrote "I have found what appears to be indisputable evidence that the appearance, and even the apparent position, of an absorption band can be profoundly modified by the presence in the absorbing vapour of a chemically inert gas" and it is now generally understood that the presence of a background gas can produce a broadening and shift of the individual spectral lines, and also an asymmetry in the line shape which becomes more marked as the pressure of the background gas increases. Over the last eighty years there has been an enormous number of studies of this phenomenon (see Fuhr *et al. NBS spec. Publ.* **366**; 1972 and Suppl. 1974, 1975, 1978). In many cases, the emitter-perturber interaction is well known, and a study of line shapes is used to determine the density of perturbing particles in the emitting region; a typical example is that of hydrogen lines perturbed by protons and electrons. However, when atomic spectra are perturbed by neutral atoms or molecules, the line shapes are analysed to

---

Gillian Peach is in the Department of Physics and Astronomy, University College London, UK.

0028-0836/81/050349-02$01.00

provide information on the basic emitter–perturber interaction. This type of analysis can be extremely valuable, since it can give information on interatomic forces where one of the atoms involved is in an excited state. Studies of atom–atom scattering predominantly yield information on interactions between two atoms that are both in their ground states (see Massey *Electronic and Ionic Impact Phenomena* 2nd edn, Vol.3, Oxford University Press; 1971).

The problem of analysing these line shapes is a fascinating and difficult one. In general, the shape of the line profile is determined by the simultaneous interaction of the emitter with many perturbers; but under certain circumstances, however, the problem can be reduced to the interaction of just one perturber with the emitting atom. If the frequency separation, $\Delta\nu$, from the line centre is sufficiently large, the mean interaction energy required, $V = h\Delta\nu$, is predominantly produced by a single perturber making a close collision with the emitter. The contribution from more distant perturbers is negligible by comparison. Thus the analysis of the shape of the line wings can give detailed information on the inner region of the atom–atom potential. This approach has been used for many years (see Kuhn *Proc. R. Soc.* **A158**, 212; 230; 1937), and an excellent modern example of such an analysis is given by York *et al.* (*J. chem. Phys.* **63**, 1052; 1975). They are able to obtain the potential energy curves for the $X^2\Sigma$ and $A^2\Pi$ states of the systems Na–Ar, Na–Kr and Na–Xe. *Ab initio* calculations of the adiabatic potentials for such systems are difficult and costly, although a good deal of progress has been made in the last few years by adopting semi-empirical model or pseudo-potential methods (see Peach *J. Phys.* B **11**, 2107; 1978; Philippe *et al. J. Phys.* B **12**, 2493; 1979).

When the frequency separation from the line centre is small the mean interaction energy V is also small, and is produced by many perturbers making weak collisions with the emitter; the probability of a close collision is very low. This picture provides the basis of the impact approximation, a simple description of which is given by Peach (*Contemp. Phys.* **16**, 17; 1975). The impact approximation predicts a lorentzian profile for the line given by

$$I(\Delta\nu) = \frac{1}{\pi}\,\frac{w(\Delta\nu)}{(\Delta\nu - d)^2 + w^2}$$

where in this case, the half width at half maximum is $w(\Delta\nu) \equiv w$, and is independent of frequency. The shift in the position of the line is given by d. Szudy and Baylis (*J. quant. Spectrosc. Radiat. Transfer* **15**, 641; 1975; **17**, 269, 681; 1977) have developed a scalar theory of line broadening which is less restrictive than the impact approximation. The scalar nature of the theory means that it is assumed that the frequency shifts produced by the individual perturbers are additive, but allows for the fact that the perturbing collisions are of finite duration. Their theory predicts a first-order correction to the pure lorentzian profile in the core of the line, which is given by $w(\Delta\nu) = w + a\Delta\nu$, where a is a constant. The parameter a depends only on the average time of duration of the collision, the nature of the interactions between the emitter in both its initial and final levels and a perturber, and the temperature of the gas. Thus the line contains a dispersive component which leads to an asymmetry in the line core. Asymmetries in the far wings of lines occur for different reasons; in this case completely different regions of the emitter–perturber potentials are being sampled. Very recently the first quantitative measurements of asymmetries in line cores have been reported (Walkup *et al. Phys. Rev. Lett.* **45**, 986; 1980). They have studied the shape of the resonance lines of sodium perturbed by helium, neon, argon, krypton and xenon. They show conclusively that the lines contain a dispersive component as predicted by the theory, and obtain accurate values for the contant a. About the same time Kielkopf (*J. Phys.* **B13**, 3813; 1980) also published width, shift and asymmetry parameters for the sodium resonance lines perturbed by the same systems. His measurements are also for similar pressures and temperature, $P = 10$–600 torr and $T = 450$K. However, his quoted errors are somewhat larger than those given by Walkup *et al.* If a pure van der Waal's potential is used for the emitter–perturber interaction, good agreement is obtained between theory and experiment for the systems Na–Ar, Na–Kr and Na–Xe. In fact this agreement is better than actually stated in the paper of Walkup *et al.* as can be shown when the full profile given by the scalar theory is used. For the systems Na–He and Na–Ne the agreement is not good; this is readily understandable since the heavier rare gases have larger dipole polarizabilities and hence the long-range part of the potential becomes relatively more important. These measurements present an exciting new challenge to the theorists. It is clear that calculations are required which use realistic interatomic potentials; on the basis of a pure van der Waal's potential, it is not possible to obtain the different values of the asymmetry parameter that are observed for the two lines. This difference gives us direct information about the difference between the $A^2\Pi$ and $A^2\Pi$ potential curves for the sodium-rare gas systems. It seems certain that more such measurements will be made, and will be used to probe regions of interatomic potentials that are not accessible in studies of the far wings of spectral lines, or in other types of atom–atom scattering experiments. □

Current issues in conservation

from Jared M. Diamond

WHEN nature reserves are looked at with the understanding gained from studies of island biogeography then the problems of avoiding major losses of the world biota seem even more difficult than at first thought. Nature reserves can be thought of as islands in a sea of man-transformed habitats. Not only will a reserve which saves only a small part of a particular habitat start out with fewer species than a larger one, it will ultimately lose a greater fraction of its species and it will lose them faster (Terborgh *Bioscience* **24**, 715; 1974). This is best illustrated by the well documented losses of species in Panama's Barro Colorado Island Reserve within this century (Willis *Ecol. Monographs* **44**, 153; 1974). Biogeographic arguments for large reserves are reinforced by genetic ones: a population with an effective size of less than 500 gradually loses the genetic variation that underlies ability to adapt, and an effective population size of less than 500 leads to inbreeding depression (Soulé and Wilcox *Conservation Biology* Sinauer; 1980). As discussion at two recent conservation meetings* in South Africa showed, even in Kruger National Park (South Africa's largest reserve) numerous key species such as lions, ground hornbills, wild dogs, and saddle-billed storks have populations whose long-term prospect of survival is only marginal from this perspective. Biologists have long felt intuitively the importance of large reserves, but lacked the detailed arguments that have been emerging in the past decade. To many participants in the South African meetings, the detailed arguments were chilling.

The meetings revealed that biologists, park managers, and government planners have important differences in vocabulary, values, and areas of concern, and that it takes effort to bridge these communication barriers. It is clear that biologists who wish to be effective have to learn the vocabulary in which public planners express themselves, to broaden their perspective so as to be aware of social and economic constraints that concern planners, and to be specific. Part of the reason for the limited success of biologists in persuading

Jared M. Diamond is Professor of Physiology at University of California Medical School, Los Angeles.

---

*A symposium on the conservation of threatened natura habitats was held in Cape Town on September 11-12 and workshop on the same topic at Tsitsikama National Park on September 14-18, 1980.

0028-0836/81/050350-02$01.00

governments to establish large reserves may be the failure to offer specific arguments in terms of fraction of species expected to be lost, expected rates of losses, and likely identities of lost species in a reserve of a given size. More generally, planning and management should be part of one continuous process, with progressively more specific stages such as: formulation of a national plan of conservation priorities; selecting areas, sites, and shapes of reserves; zoning the reserve for different uses; and developing an integrated programme of research, monitoring, and management.

The last word, management, is initially anathema to many biologists. Yet a recurring theme at these meetings was that nature reserves cannot be left alone; they have to be managed. Many biologists assume, not even consciously, that a reserve need only be set aside; nature will 'manage itself' and preserve the ecosystem, and human intervention is precisely what one wants to avoid. The fatal flaw in this reasoning is: each reserve was formerly part of a much larger biogeographic unit, and it was this larger unit which 'managed itself', that is, species and habitats rarely disappeared in the whole unit; they might disappear locally but persisted or expanded elsewhere within the unit. When a reserve becomes isolated, it is doomed to change because of the loss of processes outside its boundaries that were necessary for its maintenance. The smaller the reserve, the further and faster it will change. The reasons for these inevitable changes are multiple: reserve boundaries often intersect paths of regular seasonal movements (for example, the game fence around Kruger intersecting the former west–east seasonal migrations of game); populations isolated in a reserve may be too small to be self-sustaining, though they were able to sustain themselves in the larger system; the landscape is a shifting mosaic of habitats, and a small reserve that temporarily loses a habitat essential for certain species loses those species permanently; and extinctions of a few species lead to extinctions of other species dependent on them, via effects called trophic cascades.

A spectacular example of the need for managing even large reserves is the disaster that resulted from a hands-off approach to Tsavo National Park, where reserve status led to an increase in the number of elephants producing a complete transformation of the habitat, and ultimately a crash of the elephant population. Biologists have to resign themselves: measures such as culling, reintroductions, transfers of animals, maintaining a habitat mosaic by controlled fires, and provision of water in arid areas must be standard tools of a reserve manager. The smaller the reserve, the more effort is required, but even Kruger requires constant human intervention to avoid losses of key species and transformations of habitats. We face a cruel dilemma: without management, we will lose species and habitats; yet, only for a few species have we the biological knowledge necessary to manage well. This is not the leisurely world of academia, where one completes one's marshalling of evidence before publishing. Here, as elsewhere in the field of conservation, one must act on very limited information, often mere guesses.

Finally, the South African picture. Measured against standards of other countries, especially less wealthy countries, South Africa has done much for conservation. It has dozens of well-run reserves, including two large ones (Kruger and Kalahari Gemsbok) that are famous throughout the world. Measured against standards of what South Africa actually needs to protect its own biological heritage, there are glaring gaps (Huntley in *Biogeography and Ecology of South Africa*, 1333 (ed. M.J.A. Werger), Junk, 1978). For two habitats that covered a large fraction of South Africa, the Highveld Grassland and the Karoo, there is currently *no* large reserve at all. The Cape lowland fynbos, one of the world's richest and most distinctive temperate-zone plant communities, is represented in only a few small reserves. Unless much larger areas can be protected soon, the scale of plant extinctions in the fynbos is likely to constitute one of the most serious biological losses on Earth in the next few decades. Finally, the Pafuri area of Kruger National Park is being considered for coking coal mining. On the one hand, Pafuri is ecologically the richest area of Kruger and contains many tropical species not found elsewhere in South Africa. On the other hand, the Pafuri coking coal reserves are sufficient for only about 10 or 15 years of exploitation, and South Africa already has coal reserves adequate for its own use for centuries and exports coal to other countries. The basic concept of a national park is an irrevocable assignment of land for protection of its resources in perpetuity. It would be ironic if South Africa, which played a leading role internationally in developing this concept, were now to take the lead in dismantling it at Pafuri.

Chlorinated dioxins and the environment

from Alastair Hay

WERE it not for the bacteria and fungi to be found in the soil, the chemical industry would be having a hard time of it. Most of the organic chemicals produced by manufacturers as herbicides and pesticides are degraded by soil microorganisms. But this is not true for all chemicals. When it comes to the degradation of the polychlorinated dibenzodioxins (PCDDs) it seems that bacteria cannot cope.

The best known of the PCDDs is the tetrachlorinated isomer 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). An unwanted by-product produced in the manufacture of 2,4,5-trichlorophenols TCDD is a stable compound. A temperature in excess of 800°C is required for its thermal degradation — but, more conveniently, TCDD can be subjected to photolytic degradation provided a hydrogen donor in the form of a suitable solvent is present.

But it is not just the stability of TCDD which causes problems; the chemical is also extremely toxic with an LD_{50} as low as $0.6\mu g\ kg^{-1}$ in the guinea pig. TCDD has also been shown to be a tetratogen and carcinogen in various animal species.

It is hardly surprising, therefore, that the discharge of 250–500g of TCDD over a populated area at the Italian town of Seveso in July 1976 should be a subject of concern. The chemical was vented directly into the atmosphere when a 2,4,5-trichlorophenol reactor overheated.

Studies performed by the US Air Force had suggested that soil bacteria are capable of degrading TCDD and that the half life of the chemical — depending on the bacteria present — was between 230 and 320 days (Young *et al.* USAFA-TR-76-18; 1976).

At Seveso however, the half life is much longer and considerably in excess of 10 years according to results presented at a recent workshop* by A. di Domenico (Istituto Superiore de Sanita (ISS), Rome). It seems that photolytic degradation of TCDD did occur at Seveso in the weeks following the discharge of the chemical from the reactor. This breakdown took place when the chemical was on surfaces exposed to sunlight. Within a couple of months, however, the TCDD had been washed from plant surfaces and percolated into the first few centimetres of topsoil, where it has remained ever since. Di Domenico claims that the soil concentrations of TCDD in the most heavily contaminated region of Seveso have changed little in the 4 years since the accident.

Further evidence of the poor biodegrad-

---

*A Workshop on the Impact of Chlorinated Dioxins and related compounds on the Environment; held at the Istituto Superiore di Sanita in Rome from 22 to 24 October, 1980. The proceedings of the workshop will be published as a book by Pergamon Press.

0028-0836/81/050351-02$01.00

ability of TCDD by microorganisms was provided by R. Hütter (Swiss Federal Institute of Technology, Zurich) and I. Camani (ISS, Rome). Both reported that they had not observed any microbial degradation of TCDD under laboratory conditions. A previous report by Matsumura and Benezet (*Environ. Health Perspect.* **5**, 253; 1973) had suggested that there were some microbial strains which would detoxify TCDD, albeit with a low efficiency.

The fact that TCDD is so resistant to breakdown by soil microorganisms could mean that other PCDD isomers are equally difficult to degrade. Unfortunately there has been practically no investigation of the persistence of these other isomers in the environment. That they occur as environmental pollutants, however, is beyond dispute. Several laboratories in Europe have reported the presence of PCDD isomers in the fly ash from municipal and industrial incinerators. Scientists at the Dow Chemical Company claim that they can also be detected in ash from household chimneys, cigarettes and even charcoal-broiled steaks (see *Nature* **281**, 619; 1979).

It is not entirely clear which compounds act as precursors of PCDDs in incinerators. According to O. Hutzinger (University of Amsterdam) the principal precursors are chlorinated phenols and polyvinyl chloride (PVC). Lignin too is a likely source, says Hutzinger, and it is conceivable that inorganic chlorine might be available in incinerators for the chlorination of dibenzodioxins. Dow scientists were the first to claim that chlorine from both inorganic and organic sources could be used in the formation of PCDDs.

Whilst there is a measure of agreement on the environmental sources of PCDD, there is a marked disagreement about the type of isomer obtained in fly ash. W. Crumment (Dow Chemicals, USA) claims that fly ash contains an equal distribution of PCDD isomers containing 4,5,6,7 and 8 chlorine atoms. The tetrachlorinated isomer 2,3,7,8-TCDD is the most toxic PCDD isomer and Dow is claiming that it is present in the ash in significant amounts. Others disagree. H.R. Buser (Swiss Federal Research Station, Wadenswiss) and R.L. Harless (Environmental Protection Agency, USA) both claim that TCDD only accounts for 1% of the PCDD content of fly ash from municipal incinerators. A. Liberti (Consiglio Nazionale delle Ricerche, Rome) reported that he had not observed an equal distribution of PCDD isomers in fly ash from municipal incinerators. According to Liberti, octachlorinated dibenzodioxin — many times less toxic than TCDD — was always present in the greatest concentrations, followed closely by heptachlorinated and hexachlorinated PCDD isomers.

Can anything be done to reduce the amount of PCDs produced in large incinerators? Apparently it can. There are two alternatives; either to reduce the amount of organic chlorinated compounds burned in incinerators, or to increase the efficiency of incineration. K. Olie (University of Amsterdam) pointed out that the PCDD content of fly ash from incinerators in Holland was 20 times higher than the content of fly ash obtained from Bratislava (Czechoslovakia) where far less PVC is incinerated. As for improvements in combustion efficiency, W. Crummett pointed out that by using extra fuel in an incinerator at the Dow Chemical Company the octachlorodibenzodioxin concentration in fly ash was reduced from 180–810 p.p.b. to 0.01–0.1 p.p.b.

High temperatures are a well recognized way of degrading PCDDs. Much less is known, however, about the biological degradation of these chemicals in mammals. TCDD is the only isomer whch has been studied in any detail and Poiger and Schlatter (*Nature* **281**, 706; 1979) have reported that the chemical is hydroxylated in the rat. Metabolism of TCDD also occurs in the dog and H. Poiger (University of Zurich) now believes that this metabolism is a means of detoxifying the chemical as the acute toxicity of the metabolite is at least 100 times lower than that of TCDD itself.

One of TCDD's more worrying properties is its ability to cause malformations in animals. H. Nau (Freie Universitat, Berlin) reported that when TCDD was administered to pregnant mice only very low concentrations of the chemical were found in the embryo and fetus between gestational days 11 and 18. TCDD concentrations in the fetus were between 0.04 and 0.14% of the maternal dose, two orders of magnitude lower than the concentration in maternal liver — the organ in which TCDD accumulation was greatest. Repeated doses of $5\,\mu g\ kg^{-1}$ on days 7–11 of pregnancy produced cleft palate in 65% of fetuses whereas $25\,\mu g\ kg^{-1}$ given as a single dose on day 7 caused cleft palate in only 16% of foetuses. There were two important conclusions to be drawn from these results said Nau. Firstly, the incidence of malformations clearly depends on the frequency of dosing, and secondly, only small amounts of TCDD are required to produce fetal abnormalities.

Detecting small amounts of TCDD in humans has always been a problem. With TCDD being lipid soluble a fat biopsy was the obvious method of choice. But biopsies have their limitations, particularly because they are uncomfortable for the donor. According to C. Rappe (University of Umea) this discomfort is no longer necessary for he has developed an assay for PCDDs using only 10 ml of blood plasma.

Workers employed in tanneries constantly handle hides whch have been treated with pentachlorophenol preservatives. Rappe examined workers recently exposed to pentachlorophenol or who had been exposed 6 months previously. The results showed that pentachlorophenol is cleared from plasma when the individual is no longer exposed to the chemical. But the picture is quite different for the PCDDs which contaminate pentachlorophenol. These are far more persistent and are only cleared from plasma at a very slow rate.

Rappe's new assay will have considerable repercussions for those concerned about the health of individuals exposed to PCDDs. It will now be possible to measure the PCDD levels of workers exposed to the chemical in industry and to set more rigorous health standards. Blood measurements will be essential for interpreting clinical symptoms following individual exposure to PCDDs. In the longer term they may even assist in assessing the risk to humans of TCDD causing malformations or cancer. □

Alastair Hay is a lecturer in the Department of Chemical Pathology at the University of Leeds.

100 years ago

King of Gorontalo

Sultan of Sulu

Nature Vol. 289 29 January 1981

REVIEW ARTICLE

The origin of human cancers

John Cairns*

Imperial Cancer Research Fund, Burtonhole Lane, London NW7 1AD, UK

The limited evidence available suggests that most human cancers are not caused by conventional mutagens but are more likely to be the result of genetic transpositions. Although the molecular biology of transposition is starting to be understood, the external factors that influence its frequency have not yet been studied in any detail.

DIFFERENT human populations tend to suffer from different kinds of cancer: the inhabitants of third world countries have an excess of liver cancer, those of western nations have an excess of cancer of the breast and colon, the Japanese have an excess of stomach cancer, and so on. Most of this variation must be due to varied diets, customs and environment rather than to differences in genetic constitution, because nations have been observed to undergo changes in cancer incidence from one generation to the next and migrant populations tend to take on the pattern of cancer that is characteristic of their new homes. Cancer is therefore thought to be a preventable disease.

In the western world the two commonest cancers are cancer of the skin and the lung, and each of these happens to be so strongly dependent on a single factor (that is, sunlight and smoking) that their cause could be identified without the need for any understanding of the underlying mechanism of carcinogenesis. The preventable causes of the other common cancers are less clear-cut and it may be difficult to identify them until much more is known about mechanism. Eventually the techniques of nucleic acid chemistry should allow us to itemize all the differences in nucleotide sequence and gene expression that distinguish a cancer cell from its normal counterpart, and perhaps at that point the steps involved in carcinogenesis will cease to be in doubt. But until then we have to be content with circumstantial evidence. This review discusses the evidence and makes certain deductions about the molecular biology of human cancer.

Experimental carcinogenesis

Much of cancer research is based on the reasonable premise that we should be able to learn about human cancer by studying carcinogenesis in animals. During the 65 years since coal tar was shown to be carcinogenic for the skin of rabbits, many ways have been found for producing cancer in experimental animals. The problem is to decide which of these very diverse forms of carcinogenesis is likely to be the best model for most human cancers.

The easiest and most widely studied method is to expose the chosen target in the animal to physical or chemical agents that damage DNA. Because the carcinogenic potency of such agents is fairly well correlated with their ability to cause point mutations[1-3], these forms of cancer are commonly assumed to arise as the result of local changes in DNA sequence (in the rest of this article, the words 'mutation' and 'mutagen' are used specifically for such local changes). However, the process cannot be quite as simple as that. In most cases, repeated doses of the mutagen have to be administered over an appreciable fraction of the animal's lifespan, and in those few situations where a single dose of mutagen is carcinogenic the resulting cancer usually does not appear for a long time[4]. Thus the creation of a cancer cell is thought to involve a sequence of events of which perhaps only the early steps bear any direct relation to the interaction between mutagen and DNA. This is borne out by the observation that the later events can be caused by other agents ('promoters') that are not themselves mutagenic. The conventional explanation is that 'initiation' represents the production of mutants and that the subsequent steps of promotion are required because the relevant mutations either happen to be recessive (and so will only be detectable following deletion or somatic recombination) or usually involve repressor genes (and so are not expressed until the existing stock of repressor molecules has been depleted by further cell multiplication). But this type of interpretation is not without difficulties. The first step in the sequence seems to be too efficient[5-7]; in certain situations, even quite low doses of mutagen are capable of initiating the process in almost every exposed cell. Conversely, the later steps seem to be too sluggish[8,9]; in several instances, more than 10 cell divisions may be needed following initiation to achieve expression of the new phenotype in any of the descendent cells.

Cancers can arise following infection with certain viruses. Here the explanation seems at first sight more straightforward, for it is easy to see how novel phenotypes could result from an interaction between the genes of a virus and those of the host. However, the interval between initial infection by the virus and final appearance of the cancer can be a large fraction of the animal's lifespan[10], and so even here the exact biology of the carcinogenic process may sometimes be quite as obscure as in most chemical carcinogenesis.

Cancers can also be produced in animals by transplanting certain tissues, such as ovary[11] or embryo[12], into special sites where the cells can multiply without restraint; similarly, cancers may arise in the tissues next to implanted sheets of plastic, apparently because the cells have lost the restraining influence provided by contact with each other[13]. It seems unlikely that any conventional form of mutation can play a part in these types of carcinogenesis, particularly as the cancer cells produced by transplantation can sometimes be restored to a normal, regulated pattern of growth by transplantation back to more demanding sites[14-16].

Lastly, a number of cancers have been observed to arise spontaneously in animals that are given unlimited food, but not

---

* Present address: Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115.

0028-0836/81/050353—05$01.00

in animals whose intake of calories has been slightly restricted[17]. The mechanism of carcinogenesis by these few extra calories is completely obscure.

Those seem to be the main ways of causing cancer in animals. It is not obvious what, if anything, they have in common, and in the absence of any additional source of information we could therefore have no way of knowing which of these experimental cancers are the best models of the common human cancers; and if we do not know that, we cannot guess what class of environmental agent is likely to be the most important factor in causing human cancer.

The genetics of susceptibility to cancer

When the precise chemistry of any biological process is in doubt, it has usually been helpful to determine what kinds of mutation affect the process. This stratagem has been applied successfully in working out various metabolic pathways and in identifying many of the components involved in the synthesis of macromolecules, and it would seem to be the surest way of finding out what classes of biochemical event tend to be rate-limiting for human carcinogenesis. The method is, however, subject to certain logical restraints.

(1) When the main effect of a mutation of a multicellular animal is to alter the phenotype of one particular class of cells, the fact that it also raises the incidence of cancer in these cells may not be relevant when we come to consider the cause of other cancers. For example, some forms of inherited immunodeficiency are associated with an increased incidence of certain rare cancers arising in cells of the immune system[18]; obviously, this tells us little or nothing about the origin of the common cancers in other tissues.

(2) In general, mutations that increase the incidence of any cancer should be interpreted with some caution. Going back to the previous example, it is conceivable that the leukaemias and lymphomas of immunodeficient children arise by some special pathway, so that they may not even be a good model for the leukaemias and lymphomas that occur in normal subjects.

(3) The most informative mutations are usually those that stop things from happening. For example, mutations that block the induction of aryl hydrocarbon hydroxylase make mice more sensitive to the short-term toxicity and less sensitive to the long-term carcinogenicity of certain polycyclic aromatic hydrocarbons[19,20]. If we could show that similar mutations in humans lowered the incidence of any of the common cancers, this would suggest that these compounds are important in human carcinogenesis. So far, however, such studies in human populations have been inconclusive[21,22].

(4) In practice, much of our information may have to come from negative evidence. For instance, if a mutation that completely blocked the metabolism of certain potential carcinogens had no effect on the incidence of cancer, this would constitute very strong evidence that these chemicals and their metabolic products were probably not important. In fact, the first case we shall consider is an example of negative evidence.

If the conventional dogma is correct and cancer is usually initiated by localized somatic mutations produced by the chemical mutagens in our environment, we would expect inherited defects in DNA repair to have a marked effect on the incidence of the common cancers. Yet this turns out not to be true. The test case is the disease called xeroderma pigmentosum (XP). This arises when both copies of any one of the genes involved in the main pathway for excision repair contain an inherited defect[23,24]. The result seems to be equivalent to the effect of the *uvr* mutations of *Escherichia coli* and *Salmonella typhimurium*, which make these bacteria into exquisitely sensitive tester strains for most chemical carcinogens. Like such bacteria, the cells of XP patients have a defect in DNA excision that makes them up to 10 times more sensitive to the lethal[25-32] and mutagenic[33-38] effects of most mutagens (the main exceptions being X rays and certain simple alkylating agents). As a result of this defect, XP patients are extremely sensitive to UV light. They suffer an age-specific incidence of skin cancer that is several thousand times higher than normal and they frequently die from malignant melanoma; as these skin cancers arise on the exposed parts of the body, they are presumably the result of mutations induced by UV light in the basal cells of the skin, perhaps coupled with some radiation damage to the local immune system of the skin[39]. Yet, even though the repair defect of XP patients extends to all the tissues that have been tested and apparently affects the response to all bulky lesions and major distortions of the double helix[24], it does not result in any obvious increase in the common fatal cancers (such as cancer of the lung, alimentary canal and breast). This is a striking anomaly, but it seems to have been pointed out only once before[40]. For example, not one death from internal cancer was recorded in any of the published clinical reports on series of XP patients[41-50] summarized in Table 1; and a survey of the entire literature on XP for the past 40 years[51] (which included a search through 6 years of US death certificates[52] and therefore must comprise many hundreds of cases) has revealed only seven internal neoplasms and only two or three deaths from cancers that could not have been caused by sunlight. Now, it is possible at this point to indulge in some special pleading—by arguing, for example, that XP patients do not live long enough to experience the common internal cancers which, for some as yet unknown reason, cannot be brought forward in time as readily as skin cancers. But the most reasonable conclusion is that the commonly fatal cancers are not caused by the kind of lesions of DNA that XP patients (and bacterial *uvr* mutants) are specifically unable to repair.

That conclusion is strengthened when we consider Bloom's syndrome; an inherited disease that apparently does increase the incidence of the common cancers. Its molecular biology is not understood, but the disease is associated with a high frequency of chromosomal aberrations and exchanges[53,54]. (Interestingly, the patients' cells do not show any very marked increase in sensitivity to various mutagens[24].) From a registry of cases it has been possible to calculate very crude estimates of 'life-tables' and of relative risk for carcinomas and for leukaemias and lymphomas (Fig. 1, Table 1). We see that the death rate from cancer is raised about 100-fold in each age group (Table 1). This demonstrates that it is possible to observe such increases in risk even when the number of subjects is small and their lifespan very limited; for XP, where the numbers are greater and lifespan is somewhat longer, it should have been even easier; in this sense, therefore, Bloom's syndrome serves as a control. (Two other inherited diseases, Fanconi's anaemia and ataxia telangiectasia, are associated with chromosomal instability, and they too show a high incidence of leukaemia; however, the published reports are not sufficiently detailed for the calculation of life tables.)

Comparison of the clinical records for XP and Bloom's syndrome therefore leads to the following conclusions. (1) For most people the main source of mutagenic lesions in DNA is UV light. (2) Chemical mutagens (of the kind detected in the usual tests for mutagenicity) seem unlikely to be the rate-limiting components in most human carcinogenesis. (3) Large-scale changes in the genome (such as rearrangements and deletions) seem to be more hazardous than the local changes produced by conventional mutagens.

The karyotype of human cancers

If genetic rearrangement or deletion is a critical step in the formation of most human cancers, the change might often be large enough to produce a visible alteration in chromosomal anatomy. Unfortunately, it is not easy to obtain good metaphase spreads for the cancer cells in solid tumours, and so most studies of karyotype have been confined to the leukaemias[55] and to a few solid tumours, such as meningiomas[56] and gliomas[57], which contain neoplastic cells that grow readily *in vitro*. In these diseases, visible chromosomal rearrangements have proved to be more the rule than the exception, and for many cancers the karyotype has actually been observed to become progressively abnormal with time. It is therefore clear that at least the later stages in development of cancers are usually associated with large-scale changes in the genome.

For obvious reasons, much less is known about the earlier stages of carcinogenesis. About 90% of chronic myeloid leukaemias show a translocation from chromosome 22 to 9, which creates the characteristically shortened 22 known as the Philadelphia chromosome[58]; in several instances[59-61] the translocation has been observed to be present quite early in the sequence of events leading to the leukaemia, suggesting that it was primarily this alteration that allowed the precancerous stem cell to multiply without restraint and gradually colonize most of the bone marrow. (The exceptional cases of chronic myeloid leukaemia that do not have the translocation run a distinctly less favourable clinical course[62], and so they can probably be classified as a different kind of cancer, brought on by a different sequence of events.) Similarly, patients with ataxia telangiectasia are often found to be harbouring an expanding clone of lymphocytes with a translocation affecting chromosome 14 (refs 63, 64) and the lymphatic leukaemia which commonly occurs in these patients apparently develops within these clones. (Interestingly, myelomas and one class of lymphomas may also show translocations to chromosome 14 (ref. 65).) Those are the clearest examples where a particular change in karyotype appears to be an essential, early step in carcinogenesis. Of course, it is possible that still earlier in each of those sequences some crucial step occurred that made inevitable the subsequent emergence of cells bearing these particular rearrangements. For example, benign meningiomas have often been observed to be diploid in their early stages[66], and it is only later that they tend to be dominated by a characteristic cell that has lost one chromosome 22 (ref. 56); and there are certain kinds of cancer, such as acute myeloid leukaemia, that sometimes maintain an apparently normal karyotype throughout the whole course of the disease[66].

These studies of karyotype can therefore be interpreted in one of two ways. The obvious conclusion is that large chromosomal rearrangements are usually the crucial steps in carcinogenesis, but that there are exceptions and in certain classes of cell the critical changes either are not rearrangements or are on too small a scale to be detected by chromosomal banding techniques. It is, however, possible to argue that the cellular environment within any growing cancer is very abnormal and applies strong selection pressure for very unusual phenotypes, and these are most easily generated by large-scale rearrangements of the genome; the observed changes in karyotype could therefore be trivial secondary events that occur after all the rate-limiting steps of carcinogenesis have been completed.

The evolutionary analogy is worth pursuing a little further. The normal architecture and programmes of cell renewal in multicellular animals seem to be so arranged that the various multiplying stem cells do not, as a rule, compete with each other for territory[67]. For example, it is clear from studies of X-chromosome mosaicism that adult human tissues are made up of a large number of very small families of cells which therefore cannot be undergoing clonal selection[68,69]; similarly, the haematopoietic stem cells that circulate in the blood of mice are capable of establishing themselves in the marrow and forming clones of descendants, but they will only do so if the resident stem cells occupying the available sites in the marrow have been killed[70]. So one of the crucial steps in the development of a cancer must be the change that allows a stem cell to compete favourably with its neighbours and form an expanding clone within which further natural selection can then take place. We should therefore consider the possibility that the formation of a cancer proceeds by the same mechanism as the evolution of a new species. In other words, we might expect normal non-cancerous cells to be like species that are not evolving quickly: they will accumulate local alterations in base sequence and amino acid sequence due to forced or unforced errors in DNA replication, but these will tend to be neutral and have little effect on phenotype[71]. In contrast, a family of cancer cells should be like a species that is undergoing a rapid evolution in morphology. Such evolution is now thought to be due mainly to pleiotropic changes in the regulation of gene expression that are the result of large-scale alterations in gene arrangement and chromosome constitution[72,73]. In short, the analogy with evolution would predict that cancer should usually be due to gene rearrangements rather than to the local changes in sequence produced by conventional mutagens. Before considering what is known about the molecular biology of such rearrangements, it is worth examining the other half of the proposition, namely that carcinogenesis by conventional mutagenesis is not making a major contribution to the national incidence of cancer.

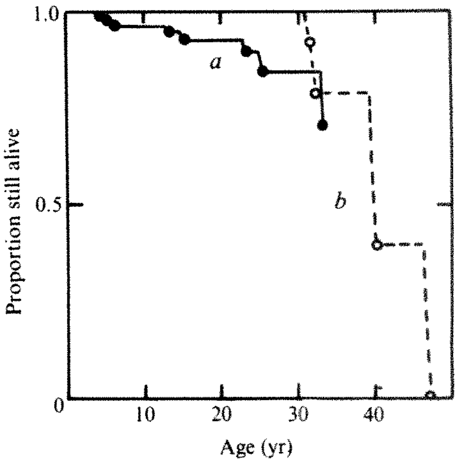


Fig. 1 Life tables for patients with Bloom's syndrome. The calculations have been based on a collection of 89 cases (J. German, unpublished), 12 of whom have died of cancer. *a*, Proportion who have not yet died of leukaemia or lymphoma; *b*, proportion who have not yet died of carcinoma.

Mutagenesis and cancer

It seems clearly established that localized lesions in DNA can be carcinogenic: for example, pyrimidine dimers must be the cause of the skin cancers produced when animals are irradiated with UV light, because the cancers can be prevented by photoreversal with visible light (a reaction that seems to be specific for dimers)[74]; similarly, the formation of O^6-alkylguanine in DNA is apparently the basis for carcinogenesis by certain alkylating agents because the cancers tend to be confined to the tissues that are unable to repair that particular lesion[75,76]. So the issue is not simply whether conventional mutagens are capable of causing cancer, but whether they reach our cells in sufficient quantity to make a major contribution to our present national incidence of cancer. In short, it is a question of the balance between dose of mutagens and the efficiency of our various pathways for DNA repair.

Although skin cancer due to UV light is the commonest cancer in western nations, it is seldom fatal and accounts for only 1% of all cancer deaths. But when there is a defect in the relevant DNA repair pathway (as in XP) the age-specific incidence of skin cancer can be raised several thousandfold and life-expectancy drop to about 20; this shows the extent to which an efficient system of repair is protecting us against one of the major hazards of our environment. Similarly, the repair pathway that handles the lesions produced by X-ray irradiation seems to be equally effective, because only about 1% of all cancer deaths are believed to be attributable to the natural sources of ionizing radiation[77]. Presumably there has always been strong selection pressure for adequate methods of DNA repair, and so it is hardly surprising to find such low levels of risk from agents like UV light and X rays, which have not changed in intensity for millions of years.

There are reasons for thinking that our exposure to the chemical mutagens in our environment is now actually somewhat less than it used to be. The argument runs as follows[78]. Experimental animals tend to show a rather limited range of responses when mutagens are added to their diet; a few, very reactive chemicals produce cancer of the oesophagus and stomach, but most mutagens produce cancer of the liver. In humans, however, liver cancer accounts for only 1% of the cancer deaths in countries like the United States and Europe; oesophageal cancer is not very common outside certain parts of Africa, Iran and China; and stomach cancer is on the decline in all industrialized nations. So it seems likely that when affluence brings us the opportunity to choose what we eat, we tend to avoid foods heavily contaminated with mutagens and are therefore able to stay well within the capabilities of our DNA repair pathways. Actually, even in undeveloped countries the level of exposure to chemical carcinogens seems to be within most people's capacity for repair; in Africa, for example, the liver cancer caused by aflatoxin is largely confined to people chronically infected with hepatitis B virus[79], which indicates that although aflatoxin is known to be one of the most potent carcinogens for certain experimental animals it is not present in a high enough concentration even in Africa to produce many cancers in normal people.

So much for the 'natural' mutagens in our environment. Fortunately, the novel unnatural mutagens that we have met in the first half of this century as a result of our industrialization seem to be no more powerful than the natural ones[80]. Occupational cancers probably account for less than 5% of all cancer deaths[81]; and the number attributable to mutagenesis is actually less than this because a large proportion of today's occupational cancers are due to asbestos, which is negligibly mutagenic[82] although it is known to produce chromosomal abnormalities[83]. These numbers are, of course, trivial compared with the effect of tobacco. This has proved to be the outstanding carcinogen of the twentieth century and is now responsible for about a third of all cancer deaths. But it is not at all clear that these cancers can be classified as due to mutagenesis. Although cigarette smoke condensate contains mutagens[84], it is much more powerful as a promoter than an initiator in experimental carcinogenesis[85,86]. Furthermore, if tobacco were mainly acting as a mutagen, there should have been some reports of lung cancer in XP patients; after all, strong sunlight (which in normal people is about as carcinogenic as heavy smoking) can produce skin cancer in these patients by the time they are 7 years old, and so smoking should have given some of them lung cancer by the time they are 30.

Table 1 Deaths from internal cancers

	Bloom's syndrome*			Xeroderma pigmentosum‡	
		obs./exp.†			obs./exp.
Age	Patient years	Leukaemias and lymphomas	Carcinomas	Patient years	All cancers (excl. skin)
0–4	340	1/0.013	0/0.017	640	0/0.059
5–14	640	3/0.020	0/0.019	886	0/0.055
15–24	327	2/0.011	0/0.016	450	0/0.036
25–34	108	2/0.005	2/0.013	213	0/0.035
35–44	16		1/0.008	109	0/0.038
45–54	3		1/0.005	54	0/0.107
Totals		12/0.127			0/0.330

The observed values were obtained from various clinical reports. For Bloom's syndrome the estimates are probably fairly reliable, because the condition is very rare and so the fate of each case is usually followed rather closely; of these 89 cases, 17 are known to have died (12 from cancer). Because XP is much less rare, cases tend to be reported but not followed up, and so the existing literature may be used to give some idea of their incidence of cancer but will give an underestimate of their death rate. During the limited period when these 140 cases were observed, 71 had had skin cancer (which was fatal in 8 cases), 1 had had a brain tumour and 1 had chronic myeloid leukaemia.

*89 cases (J. German, unpublished).

†Expected values were calculated from the current age-specific cancer death rates in England and Wales.

‡140 cases reported in refs 41–50.

All these arguments are indirect, but they do suggest rather strongly that local changes in DNA sequence, produced by conventional mutagens, are probably making only a minor contribution to our national death rates.

Genetic transposition as a cause of cancer

In the last decade, the routes leading to genetic diversity have become much better understood[87,88]. Diffusion of heritable information is now known to occur not only by 'legitimate' recombination between homologous regions of different genomes, but also by 'illegitimate' recombination between largely nonhomologous regions (which can be in different parts of the genome or even in different individuals). The second of these processes is equivalent to a chromosomal rearrangement and allows the shuffling of certain modules of information ('transposons') so that new combinations of genes can be tested for survival advantage. The process depends on the presence of certain short sequences in the DNA that are the substrates for various recombinational enzymes ('transposases'), and it represents a much more drastic form of genetic variation than localized changes in base sequence because it can alter the expression of whole regions of the genome.

Transpositions can be site specific (for example, the changes in surface antigens of *Salmonella* and certain trypanosomes) or nonspecific (such as the spread of antibiotic resistance between different bacteria); they are usually solitary, but they can occur in clusters; and their frequency can vary from 1 per cell division down to 10^{-10}. A transposon can be the substrate for a transposase encoded elsewhere in the genome or it can make its own specific transposase. As the result of moving to a new site, a transposon can start or stop being transcribed; but the main effect can be not on itself so much as on neighbouring genes, which can be turned on or off or made unstable by its presence.

Transposition is therefore a very special form of genetic regulation because it can be programmed to occur at any predetermined rate, can be concerned with any desired number of alternative states, and can be made virtually irreversible. For example, the switch in the mating type of yeasts is the result of a transposition which occurs at a rate that is itself under separate genetic control[89]; but the change has all the characteristics of a step in differentiation because it concerns the function of a large number of genes but normally is confined to one particular branch of the cell lineage (the mother cells, rather than the buds)[90]. It is not yet clear whether transposition is important in vertebrate development. As the result of certain experiments on nuclear transplantation, there has been a tendency in recent years to regard the nuclei of the somatic cells of vertebrates as totipotential and therefore not fundamentally different from the germ line[91]. However, the fact remains that although it is easy to produce a normal animal using the nucleus taken from a fertilized egg or an early blastula cell, no one has ever succeeded in producing a normal animal using the nucleus taken from a somatic cell of an adult. It is therefore quite conceivable that changes in gene arrangement do occur during vertebrate development, especially as it is now known that antibody diversity is generated by specific transpositions that occur during the differentiation of B lymphocytes[92,93].

As pointed out in the earlier sections of this review, there are good reasons for believing that the steps leading to human cancer are more likely to be transpositions than localized changes in sequence. It is therefore important to find out what classes of external agents or features of cellular behaviour raise the frequency of such transpositions. Interestingly, conventional mutagens often appear to have no effect. Indeed, it was the inability of any mutagen to raise the spontaneous reversion rate of certain forward mutations in *E. coli* that initially marked out these mutations as being unusual[94] and eventually led to the discovery that they were due to the movement of insertion sequences; similarly, the movement of certain transposable elements in yeast has been found to occur 'spontaneously' but does not appear to be strongly catalysed by any of the common mutagens[95]. However, rates of transposition can be influenced by external factors. For example, transpositions are only a minor cause of mutation in rapidly multiplying bacteria[96], but they are a major source of forward mutations in stationary cultures[97]; the growth of *Drosophila* cells *in vitro* leads to the proliferation of certain repetitive sequences[98]; the frequency of the transpositions that produce variegation in *Drosophila* and in *Antirrhinum* petals is inversely related to temperature[99,100]; and the integration of a tumour virus can be associated with transpositions in the neighbouring host DNA[101]. Lastly, many mutagens are known to cause sister chromatid exchanges and other chromosomal interactions in mammalian cells, although it is not always clear to what extent these gross changes are functionally equivalent to genetic transpositions.

The assay of mutagens using bacterial tester strains has been one of the most fruitful practical results of molecular genetics. Its success stems from the fact that the chemistry of DNA is the same for all forms of life and so the chemistry of mutagenesis (and of repair) also tends to be the same. The molecular biology of transposition promises to be more idiosyncratic. The transposable genetic elements were discovered by McClintock[102] in 1950 and their possible role in carcinogenesis was first discussed by Temin[103] in 1970, but it is still too early for there to have been much systematic molecular biological investigation of the factors that trigger transpositions, either in prokaryotes or in eukaryotes. The rate of movement of most transposons will have been subject to evolutionary pressures and many transposons will therefore have acquired their own individual controls. Thus it may not be a simple matter to devise a general assay for the factors that drive the carcinogenic transpositions.

I thank Errol Friedberg, Richard Peto, Jeffrey Miller, Julian Gross and Bryn Bridges for useful advice and discussion, James German and Kenneth Kraemer for unpublished information about patients with Bloom's syndrome and XP, and David Roberts for a literature search for clinical reports on XP.

1. Meselson, M. & Russell, K. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 1473–1481 (Cold Spring Harbor Laboratory, New York, 1977).
2. Ames, B. N. & Hooper, K. *Nature* **274**, 19–20 (1978).
3. Bartsch, H. *et al.* in *Molecular and Cellular Aspects of Carcinogen Screening Tests* (eds Montesano, R., Bartsch, H. & Tomatis, L.) 179–241 (IARC, Lyon, 1980).
4. Laerum, O. D. & Rajewsky, M. F. *J. natn. Cancer Inst.* **55**, 1177–1187 (1975).
5. Mondal, S. & Heidelberger, C. *Proc. natn. Acad. Sci. U.S.A.* **65**, 219–225 (1970).
6. Parodi, S. & Brambilla, G. *Mutat. Res.* **47**, 53–74 (1977).
7. Kennedy, A. R., Fox, M., Murphy, G. & Little, J. B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1334–1338 (1978).
9. Barrett, J. C. & Ts'O, P. O. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3761–3765 (1978).
10. Allison, A. C., Chesterman, F. C. & Baron, S. *J. natn. Cancer Inst.* **38**, 567–572 (1967).
11. Furth, J. *J. natn. Cancer Inst.* **8**, 7–16 (1947).
12. Stevens, L. C. *Devl Biol.* **21**, 364–382 (1970).
13. Karp, R. D. *et al. J. natn. Cancer Inst.* **51**, 1275–1285 (1973).
14. Brinster, R. L. *J. exp. Med.* **140**, 1049–1056 (1974).
15. Mintz, B. & Illmensee, K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3585–3589 (1975).
16. Papaioannou, V. E., McBurney, M. W., Gardner, R. L. & Evans, M. J. *Nature* **258**, 70–73 (1975).
17. Tannenbaum, A. & Silverstone, H. *Adv. Cancer Res.* **1**, 451–501 (1953).
18. Good, R. A. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1026–1032 (1972).
19. Kouri, R. E., Ratrie, H. & Whitmire, C. E. *J. natn. Cancer Inst.* **51**, 197–200 (1973).
20. Nebert, D. W. *J. natn. Cancer Inst.* **64**, 1279–1290 (1980).
21. Atlas, S. A., Vesell, E. S. & Nebert, D. W. *Cancer Res.* **36**, 4619–4630 (1976).
22. Paigen, B. *et al. Cancer Res.* **37**, 1829–1837 (1977).
23. Cleaver, J. E. & Bootsma, D. *A. Rev. Genet.* **9**, 19–38 (1975).
24. Friedberg, E. C., Ehmann, U. K. & Williams, J. I. *Adv. Radiat. Biol.* **8**, 85–174 (1979).
25. Kleijer, W. J., Lohman, P. H. M., Mulder, M. P. & Bootsma, D. *Mutat. Res.* **9**, 517–523 (1970).
26. Cleaver, J. E. *Mutat. Res.* **12**, 453–462 (1971).
27. Slor, H. *Mutat. Res.* **19**, 231–235 (1973).
28. Maher, V. M., Birch, N., Otto, J. R. & McCormick, J. J. *J. natn. Cancer Inst.* **54**, 1287–1294 (1975).
29. Maher, V. M., Ouellette, L. M., Mittlestat, M. & McCormick, J. J. *Nature* **258**, 760–763 (1975).
30. Fornace, A. J. & Kohn, K. W. *Biochim. biophys. Acta* **435**, 95–103 (1976).
31. Day, R. S., Scudiero, D. & Dimattina, M. *Mutat. Res.* **50**, 383–394 (1978).
32. Andrews, A. D., Barrett, S. F. & Robbins, J. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1984–1988 (1978).
33. Maher, V. M., Ouellette, L. M., Curren, R. D. & McCormick, J. J. *Nature* **261**, 593–595 (1976).
34. Maher, V. M., McCormick, J. J., Grover, P. L. & Sims, P. *Mutat. Res.* **43**, 117–138 (1977).
35. Arlett, C. F. & Harcourt, S. A. in *DNA Repair Mechanisms* (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 633–636 (Academic, New York, 1978).
36. Maher, V. M., Dorney, D. J., Mendrala, A. L., Konze-Thomas, B. & McCormick, J. J. *Mutat. Res.* **62**, 311–323 (1979).
37. Myhr, B. C., Turnbull, D. & DiPaolo, J. A. *Mutat Res.* **62**, 341–353 (1979).
38. Glover, T. W., Chang, C. C., Trosko, J. E. & Li, S. S. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3982–3986 (1979).
39. Bridges, B. & Strauss, G. *Nature* **283**, 523–524 (1980).
40. German, J. in *Proc. 6th int. Congr. Radiation Research* (eds Okada, S. *et al.*) 496–505 (Tokyo University Press, 1979).
41. Reese, A. B. & Wilber, I. E. *Am. J. Ophthal.* **26**, 901–911 (1943).
42. Berlin, C. & Tager, A. *Dermatologica* **116**, 27–35 (1958).
43. El-Hefnawi, H., El-Nabawi, M. & Rasheed, A. *Br. J. Derm.* **74**, 201–213 (1962).
44. El-Hefnawi, H. & Rasheed, A. *J. Egypt. med. Ass.* **46**, 985–1006 (1963); **49**, 419–434 (1966).
45. Lemaire, M. & Gaumond, E. *Can. med. Ass. J.* **92**, 406–412 (1965).
46. Siegelman, M. H. & Sutow, W. W. *J. Pediat.* **67**, 625–630 (1965).
47. Sambasivan, T. S., Siddappa, K. & Indrakumar, S. V. *Indian J. Derm.* **11**, 79–81 (1966).
48. Robbins, J. H., Kraemer, K. H., Lutzner, M. A. Festoff, B. W. & Coon, H. G. *Ann. intern. Med.* **80**, 221–248 (1974).
49. Takebe, H. *et al. Cancer Res.* **37**, 490–495 (1977).
50. Pawsey, S. A., Magnus, I. A., Ramsay, C. A., Benson, P. F. & Giannelli, F. *Q. Jl Med. new Ser.* **48**, 179–210 (1979).
51. Kraemer, K. H. in *Clinical Dermatology* (eds Demis, D. J., Dobson, R. L. & McGuire, J.) (Harper & Row, New York, in the press).
52. Miller, R. W. *Pediat. Res.* **3**, 389–397 (1969).
53. German, J., Archibald, R. & Bloom, D. *Science* **148**, 506–507 (1965).
54. German, J. in *Chromosomes and Cancer* (ed. German, J.) 601–617 (Wiley, New York, 1974).
55. Mitelman, F. & Levan, G. *Cancer Genet. Cytogenet.* **1**, 29–32 (1979).
56. Zankl, H. Zang, K. D. *Cancer Genet. Cytogenet.* **1**, 351–356 (1980).
57. Mark, J., Westermark, B., Pontén, J. & Hugosson, R. *Hereditas* **87**, 243–260 (1977).
58. Rowley, J. D. *Nature* **243**, 290–293 (1973).
59. Canellos, G. P. & Whang-Peng, J. *Lancet* **ii**, 1227–1229 (1972).
60. Baccarani, M., Zaccaria, A. & Tura, S. *Lancet* **ii**, 1094 (1973).
61. Verhest, A. & van Schoubroeck, F. *Lancet* **ii**, 1386 (1973).
62. Whang-Peng, J., Canellos, G. P., Carbone, P. P. & Tijo, J. H. *Blood* **32**, 755–766 (1968).
63. Cohen, M. M., Shaham, M., Dagan, J., Shmueli, E. & Kohn, G. *Cytogenet. Cell Genet.* **15**, 338–356 (1975).
64. McCaw, B. K., Hecht, F., Harnden, D. G. & Teplitz, R. L. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2071–2075 (1975).
65. Mark, J., Dahlenfors, R. & Ekedahl, C. *Cancer Genet. Cytogenet.* **1**, 39–56 (1979).
66. Mark, J. *Adv. Cancer Res.* **24**, 165–222 (1977).
67. Cairns, J. *Nature* **225**, 197–200 (1975).
68. Gartler, S. M., Gandini, E., Hutchison, H. T., Campbell, B. & Zechhi, G. *Ann. hum. Genet.* **35**, 1–7 (1971).
69. Feder, N. *Nature* **263**, 67–69 (1976).
70. Micklem, H. S., Ogden, D. A., Evans, E. P., Ford, C. E. & Gray, J. G. *Cell Tissue Kinet.* **8**, 233–248 (1975).
71. Wilson, A. C., Carlson, S. S. & White, T. J. *A. Rev. Biochem.* **46**, 573–639 (1977).
72. Wilson, A. C., Bush, G. L., Case, S. M. & King, M. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5061–5065 (1975).
73. Bush, G. L., Case, S. M., Wilson, A. C. & Patton, J. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3942–3946 (1977).
74. Hart, R. W., Setlow, R. B. & Woodhead, A. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5574–5578 (1977).
75. Goth, R. & Rajewsky, M. F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 639–643 (1974).
76. Margison, G. P. & Kleihues, P. *Biochem. J.* **148**, 521–525 (1975).
77. Jablon, S. & Bailar, J. C. *Preventative Med.* **9**, 219–226 (1980).
78. Peto, R. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 1403–1428 (Cold Spring Harbor Laboratory, New York, 1977).
79. Larouze, B. *et al. Lancet* **ii**, 534–538 (1976).
80. Doll, R. *Proc. R. Soc.* **B205**, 47–61 (1979).
81. Higginson, J. & Muir, C. S. *J. natn. Cancer Inst.* **63**, 1291–1298 (1979).
82. Huang, S. L. *Mutat. Res.* **68**, 265–274 (1979).
83. Sincock, A. & Seabright, M. *Nature* **257**, 56–58 (1975).
84. Kier, L. D., Yamasaki, E. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4159–4163 (1974).
85. Roe, F. J. C., Salaman, M. H. & Cohen, J. *Br. J. Cancer* **13**, 623–633 (1959).
86. van Duuren, B. L., Sivak, A., Segal, A., Orris, L. & Langseth, L. *J. natn. Cancer Inst.* **37**, 519–526 (1966).
87. Nevers, P. & Saedler, H. *Nature* **268**, 109–115 (1977).
88. Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
89. Hicks, J. B., Strathern, J. N. & Herskowitz, I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 457–462 (Cold Spring Harbor Laboratory, New York, 1977).
90. Strathern, J. N. & Herskowitz, I. *Cell* **17**, 371–381 (1979).
91. Gurdon, J. B., Laskey, R. A. & Reeves, O. R. *J. Embryol. exp. Morph.* **34**, 93–112 (1975).
92. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa. S. *Nature* **280**, 288–294 (1979).
93. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450–3454 (1979).
94. Saedler, H. & Starlinger, P. *Molec. gen. Genet.* **100**, 178–189 (1967).
95. Roeder, G. S., Farabaugh, P. J., Chaleff, D. T. & Fink, G. R. *Science* **209**, 1375–1380 (1980).
96. Calos, M. P., Johnsrud, L. & Miller, J. H. *Cell* **13**, 411–418 (1978).
97. Arber, W. *et al. Cold Spring Harb. Symp. quant. Biol.* **43**, 1197–1208 (1978).
98. Potter, S. S., Brorein, W. J., Dunsmuir, P. & Rubin, G. M. *Cell* **17**, 415–427 (1979).
99. Gowen, J. W. & Gay, E. H. *Science* **77**, 312 (1933).
100. Harrison, B. J. & Fincham, J. R. S. *Heredity* **19**, 237–258 (1964).
101. Botchan, M., Stringer, J., Mitchison, T. & Sambrook, J. *Cell* **20**, 143–152 (1980).
102. McClintock, B. *Proc. natn. Acad. Sci. U.S.A.* **36**, 344–355 (1950).
103. Temin, H. M. *Perspectives biol. Med.* **14**, 11–26 (1970).

ARTICLES

Thermodynamic stability and kinetics of perovskite dissolution

H. Wayne Nesbitt, G. Michael Bancroft, William S. Fyfe, Suresh N. Karkhanis & Akio Nishijima

Centre for Chemical Physics and Department of Geology, University of Western Ontario, London, Canada N6A 3K7

Shigemitsu Shin

National Chemical Laboratory for Industry, Tsukuba, Japan

Perovskite, a SYNROC host mineral for nuclear wastes, is thermodynamically unstable in natural waters and in association with common minerals. Leach experiments demonstrate that $CaTiO_3$ (perovskite), $SrTiO_3$ and $BaTiO_3$ are as reactive as some silicate glasses below 100 °C, but leach much more slowly than glasses above 100 °C.

THE use of titanates and zirconates as hosts for nuclear waste in the SYNROC process has recently been proposed[1,2]. Three minerals are used as hosts—perovskite ($CaTiO_3$), Ba-hollandite ($BaAl_2Ti_6O_{16}$) and zirconolite ($CaZrTi_2O_7$). SYNROC relies heavily on geological and geochemical observations in selecting stable host minerals. Although the SYNROC minerals are not thermodynamically compatible with some siliceous rocks, the minerals are considered likely to be thermodynamically stable in the presence of water. Moreover, it has been reported that these minerals are kinetically stable in extreme hydrothermal conditions (up to 900 °C and 5,000 bar), and are thus vastly more stable to leaching than borosilicate glasses[1,2].

We report here thermodynamic calculations and leaching kinetics from 25 °C to 300 °C on the simplest of the three SYNROC minerals, perovskite, and its Sr and Ba analogues. Perovskite is suitable for such a thermodynamic and kinetic study because its dissolution reactions are simple, and the thermodynamic data for perovskite and its reaction products are known for a wide range of temperatures. By comparison with glasses, the leach rate of perovskite and its analogues is very insensitive to temperature, due to an interesting combination of the thermodynamics and kinetics of $MTiO_3$ (M = Ca, Sr, Ba) dissolution.

Thermodynamic stability

We first consider the thermodynamic stability of perovskite in the presence of the most common silicates, and then consider the three most likely reactions of perovskite in natural waters.

The plagioclase and alkali feldspars are the most abundant minerals in the Earth's crust followed by quartz (and polymorphs) and orthopyroxenes. These three mineral groups contain 50 mol % SiO_2 or more and constitute between 80 and 90% of all minerals present in the crust[3]. Perovskite stability in the presence of these siliceous minerals must be considered for they may react directly or indirectly (via a fluid) with the titanate.

Consider the reactions:

$$\underset{\text{perovskite}}{CaTiO_3} + \underset{\text{quartz}}{SiO_2(c)} \rightleftharpoons \underset{\text{sphene}}{CaTiSiO_5} \qquad (1)$$

$$\underset{\text{perovskite}}{2\,CaTiO_3} + \underset{\text{plagioclase}}{NaAlSi_3O_8} \rightleftharpoons \underset{\text{sphene}}{2\,CaTiSiO_5} + \underset{\text{nepheline}}{NaAlSiO_4} \qquad (2)$$

$$\underset{\text{perovskite}}{CaTiO_3} + \underset{\text{alkali feldspar}}{KAlSi_3O_8} \rightleftharpoons \underset{\text{sphene}}{CaTiSiO_5} + \underset{\text{leucite}}{KAlSi_2O_6} \qquad (3)$$

$$\underset{\text{perovskite}}{CaTiO_3} + \underset{\text{orthopyroxene}}{2\,MgSiO_3} \rightleftharpoons \underset{\text{sphene}}{CaTiSiO_5} + \underset{\text{olivine}}{Mg_2SiO_4} \qquad (4)$$

The thermochemical data for these phases[4] have been used to calculate the free energies for each of the reactions (ΔG^0_R) at various temperatures. The free energies of all reaction are negative at 25 °C indicating that sphene is stable relative to perovskite in the presence of quartz, feldspars and orthopyroxenes. The ΔG^0_R values become progressively more positive at higher temperatures but all values remain more negative than −10 kJ even at 900 °C. The incompatibility of perovskite and quartz, orthopyroxene and particularly feldspar severely restricts the repository or vault type if this potential for perovskite alteration is to be minimized. Virtually all plutonic and most volcanic rocks would have to be discarded for they generally contain feldspars and/or orthopyroxene. The only igneous rocks suitable would be low Si, high Ca rocks such as those given by Carmichael *et al.*[5]. Alternatively, non-siliceous rocks such as evaporite deposits and carbonates would minimize this source of potential alteration.

Kaolinite [$Al_2Si_2O_5(OH)_4$] is one of the most common clay minerals produced by weathering of aluminous silicates (such as feldspars) near the Earth's surface[6]; hence it is an important constituent of many muds, mudstones, shales and argillaceous rocks. Perovskite stability in its presence can be determined from:

$$\underset{\text{perovskite}}{CaTiO_3} + \underset{\text{kaolinite}}{\tfrac{1}{2}Al_2Si_2O_5(OH)_4} + \underset{\text{water}}{\tfrac{1}{2}H_2O} \rightleftharpoons \underset{\text{sphene}}{CaTiSiO_5} + \underset{\text{gibbsite}}{Al(OH)_3} \qquad (5)$$

The free energy of the reaction is negative at low temperature indicating that clay minerals can supply sufficient SiO_2 to destabilize perovskite. The result demonstrates that even shales and argillaceous rocks would be inadequate vaults in that they provide the potential for perovskite alteration. Ringwood *et al.*[1,2] recognized the incompatibility of quartz and perovskite and they emphasized that thermodynamically compatible geological materials should be used for vaults. The above calculations emphasize that the choice of suitable host rocks is much more restricted than they suggested—if thermodynamic stability dictates the choice.

Experimental evidence shows that perovskite does react with common minerals at measurable rates. Schuiling and Vink[7] studied the hydrothermal stability of perovskite in the presence of siliceous minerals. The results agree with the above calculations. Furthermore, they converted at least 5% of the perovskite to sphene in only 3 weeks (700 °C, 1,000 bar pressure), as this amount is required for positive detection by X-ray diffraction. Their studies demonstrate that perovskite is unstable in the presence of quartz, feldspars and orthopyroxene.

0028–0836/81/050358—05$01.00

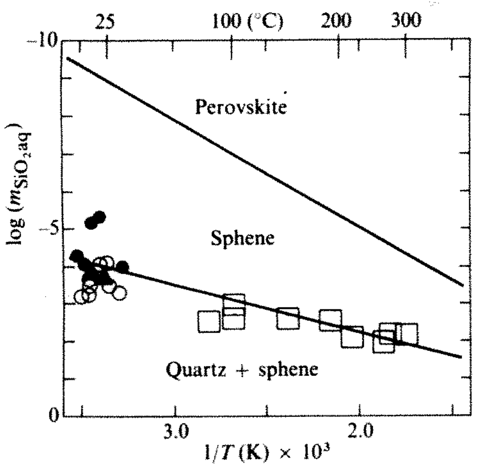


Fig. 1 log $m_{SiO_2(aq)}$ plotted against $1/T$ for reaction (6). The m_{SiO_2}. T values are plotted for groundwaters emanating for dunites and peridotites (●), serpentinites (○), limestones (△) and hydrothermal waters (□). These all plot very close to the lower solid line representing solutions saturated with quartz.

Aqueous solutions are almost ubiquitous in the crust, even to great depths[8] and of all interactions involving perovskite, reaction with subsurface waters are most likely; consequently the stability of perovskite in natural groundwaters must be investigated. Two reactions limit the stability of perovskite. One is

$$\underset{\text{perovskite}}{CaTiO_3} + \underset{\substack{\text{aqueous}\\\text{silica}}}{SiO_2(aq)} \rightleftharpoons \underset{\text{sphene}}{CaTiSiO_5} \qquad (6)$$

where

$$\log K_6 = -\log[SiO_2(aq)] \qquad (7)$$

Square brackets denote activities. The $\Delta G^0_{R(6)}$ and $\log K_6$ values were calculated and the equilibrium aqueous silica concentrations plotted against the inverse of temperature (Fig. 1). Groundwaters emanating from dunites, peridotites[9,10], serpentinites and limestones as well as hydrothermal waters have been plotted and all fall within the sphene stability field. Perovskite is unstable in these waters. Quartz solubility is given by

$$\underset{\text{quartz}}{SiO_2(c)} \rightleftharpoons \underset{\text{aqueous silica}}{SiO_2(aq)} \qquad (8)$$

where

$$\log K_8 = \log[SiO_2(aq)] \qquad (9)$$

The lower solid line (Fig. 1) represents solutions saturated with respect to quartz.

Carbon dioxide is dissolved in most natural groundwater[11] and is one of the most abundant gases dissolved in hydrothermal waters[9]. Carbon dioxide in natural waters (carbonic acid) may enhance the thermodynamic and kinetic instability of perovskite as shown below

$$\underset{\text{perovskite}}{CaTiO_3} + CO_2(g) \rightleftharpoons \underset{\text{rutile}}{TiO_2} + \underset{\text{calcite}}{CaCO_3} \qquad (10)$$

For this equilibrium relation

$$\Delta G^{\circ}_{R(10)} = -RT \ln K_{(10)} = +RT \ln P_{CO_2} \qquad (11)$$

Both the $\Delta G^{\circ}_{R(10)}$ and the equilibrium carbon dioxide pressure (P_{CO_2}) have been evaluated. The logarithm of the equilibrium partial pressure has been plotted against the reciprocal of temperature (Fig. 2). Some near-surface groundwaters from limestones[11] and from some ultramafic rocks[12] plot well within the rutile + calcite field. White *et al.*[11] compiled typical groundwaters emanating from a large variety of igneous, metamorphic and sedimentary rocks and all have P_{CO_2} values sufficiently high to alter perovskite to the much more stable (and common) assemblage, calcite + rutile. The calculations demonstrate that perovskites are generally unstable in CO_2-bearing groundwater and hydrothermal waters. Limestones and dolomites would, therefore, be unsatisfactory repository materials at least in environments where the temperatures are < 200 °C.

Table 1 Weight loss and Ca^{2+} leached into solution* for perovskite single crystal

	Conditions	Weight loss (μg)	Total Ca^{2+} (μg)	Leach rate† ($kg\ m^{-2}\ s^{-1}$)
I	300 °C, 1 week	500	430	8.9×10^{-9}
II	300 °C, 2 weeks	100	54	5.6×10^{-10}
III	100 °C, 1 week	220	94	2.2×10^{-9}

* This single crystal from Magnet Cove, Arkansas, weighed ~120 mg and had a ~0.8 cm^2 geometric surface area. The electron probe analysis of this sample showed it to be homogeneous and gave the analyses: 38.9% CaO, 42.3% TiO_2, 7.8% Nb_2O_5, 5.9% FeO, 1.2% Al_2O_3 and 0.8% SiO_2 (total 96.9%). The high Fe and Nb analyses are common for perovskites in carbonatites from Magnet Cove[13]. The remainder of the analysis is most likely rare earths[13]. A few extremely small inclusions of apatite were also seen, and analysed by the probe.

† Based on Ca^{2+} leached into solution. $kg\ m^{-2}\ sec^{-1} = 1.16 \times 10^{-4}\ g\ cm^{-2}\ day^{-1}$.

A third reaction also influences perovskite stability:

$$\underset{\text{perovskite}}{CaTiO_3} + H_2O(l) \rightleftharpoons Ca^{2+}_{(aq)} + \underset{\text{rutile}}{TiO_2} + 2\,OH^-_{(aq)} \qquad (12)$$

and when recast using hydrogen ions is

$$\underset{\text{perovskite}}{CaTiO_3} + 2\,H^+ \rightleftharpoons Ca^{2+}_{(aq)} + \underset{\text{rutile}}{TiO_2} + H_2O(l) \qquad (13)$$

Both the $\Delta G^{\circ}_{R(13)}$ and the equilibrium constant (K_{13}) have been calculated where

$$\log K_{13} = \log [Ca^{2+}]/[H^+]^2$$

Square brackets denote activities; activities of the solids and $H_2O(l)$ are taken as unity and the thermodynamic data is from Robie *et al.*[4] and Helgeson[10].

Equilibrium diagrams which include the relations of equations (6) and (13) have been constructed for 25 °C. Figure 3 delineates the stability of the three Ti-rich solids (plus amorphous silica) in the presence of aqueous solutions. Perovskite is stable only in groundwaters of very high $[Ca^{2+}]/[H^+]^2$ values ($> 10^{18.3}$) and very low aqueous silica values ($< 10^{-9}$). Solution compositions typical of near-surface ground waters in igneous rocks (ref. 12, Tables 1, 2 and 3) have been plotted (Fig. 3). Waters derived from limestones, dolostones and evaporites (data from ref. 12, Tables 6, 7 and 8) plot in the same region. As Fig. 3 shows, natural waters are too rich in $SiO_2(aq)$ and have

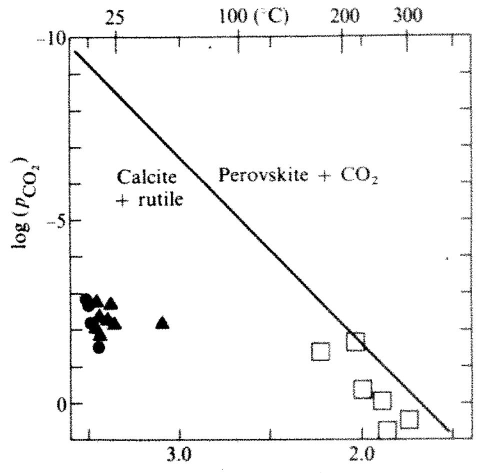


Fig. 2 log P_{CO_2} plotted against $1/T$ for reaction (10). P_{CO_2}, T values are plotted for near-surface groundwaters from limestones (△), ultramafic rocks (●) and hydrothermal solutions (□).

$[Ca^{2+}]/[H^+]^2$ values too low to be equilibrated with perovskite. Furthermore these variables are not greatly influenced by the nature of the rock from which they were collected. This is not surprising because the composition of a groundwater is the net result of reactions with all minerals and rocks it has contacted, not just the last materials encountered. For example, groundwaters which have flowed through an igneous rock may have picked up appreciable SiO_2 (aq) by dissolution of feldspar. The SiO_2(aq) will not necessarily be precipitated if the groundwater contacts limestones or evaporites. We conclude that there is probably no 'inert' vault material for perovskite because the most likely reaction is between perovskite and groundwaters, the groundwaters having contacted and reacted with various rock types.

Experimental

Four different types of samples were studied: a perovskite single crystal from Magnet Cove, Arkansas[13,14], perovskite grains from Wolgedee Hill, Western Australia, and synthetic $SrTiO_3$ and $BaTiO_3$. All samples showed excellent X-ray powder diffraction patterns characteristic of the titanate (see footnotes to the tables for other details).

Leaching experiments were performed using stainless steel autoclaves and distilled water. Because we expected very low leaching rates at low temperatures, we began our leaching experiments at 300 °C. The Magnet Cove sample was polished with 0.5 μm alumina on a micropolishing cloth before leaching for one week at 300 °C in 60 ml water, repolished, and then leached for two weeks in identical conditions. After repolishing, the sample was leached for one week at 100 °C.

The other three samples were placed in 10 ml of water in smaller stainless steel autoclaves—usually for one week. Several leach runs were made on the Australian perovskite at 100 °C, before using the same samples for leaching at 300 °C, and then 100 °C again. The samples were then ground to expose fresh surface before leaching for several weeks at 25 °C. The Ba and Sr titanates were leached at 25 °C, 100 °C, 200 °C and 300 °C often for two or three weeks (one week at a time). They were not polished or scraped between runs.

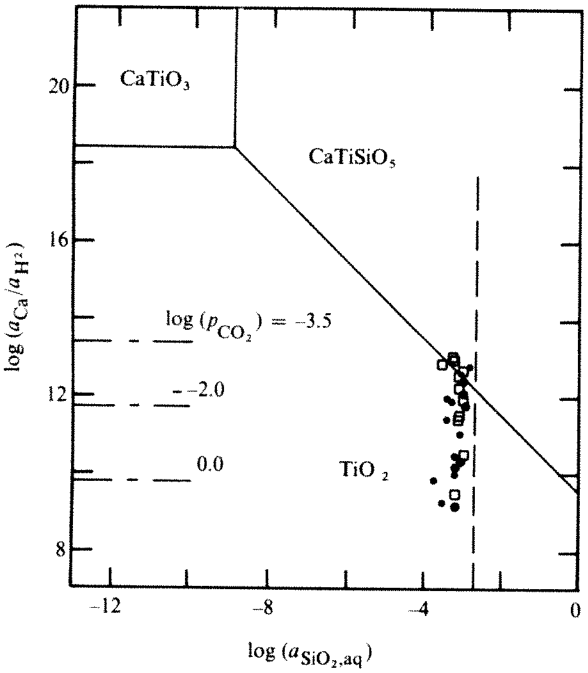


Fig. 3 $\log[a_{Ca^{2+}}/a_{H^+}{}^2]$ plotted against $\log a_{SiO_2(aq)}$ combining equations (6) and (13) at 25 °C. Solution compositions are plotted for basalts and gabbros (□), rhyolites and granites (●). The equilibrium line delineating the stability fields of sphene and rutile is obtained by subtracting the log k values of reaction (6) and (13). The dash-dot lines represent calcite solubility at the stipulated partial pressures of CO_2.

After leaching, the samples were examined by a Philips 501 SEM and an ESCA 36 spectrometer, and the solution was analysed for Ca, Sr, Ba and sometimes Ti by atomic absorption spectroscopy.

Results and discussion

In the first set of experiments with the Magnet Cove single crystal, we were able to obtain both an accurate weight loss and total amount of leached Ca^{2+} (Table 1). Although we did not deliberately exclude CO_2, the weight loss and Ca^{2+} in solution are consistent qualitatively with reaction (13) taking place. According to this equation the total weight of leached Ca^{2+} should be ~72% of the weight loss. Note that the leach rate in experiment II (5.6×10^{-10} kg m^{-2} s^{-1} or 4.8×10^{-6} g cm^{-2} day^{-1}) is much lower than that in I (8.9×10^{-9} kg m^{-2} s^{-1}) while the leach rate at 100 °C if anything is higher than the latter 300 °C leach rate. Scanning electron micrographs on the leached crystal indicated new crystals on the surface, but these were not large enough to identify.

Four separate ~0.5-g samples of the Western Australia perovskite grains were then leached for seven days in 10 ml of distilled water at 100 °C. The average amount of Ca^{2+} ion leached into solution was 78 ± 20 μg. An approximate surface area of 10 ± 2 cm^2 was obtained by microscopic examination, and the leach rate becomes 1×10^{-10} kg m^{-2} s^{-1}, substantially lower than that obtained on the single crystal. Two of the perovskite grain samples were then leached at 300 °C and then 100 °C for one week, yielding leach rates of 1×10^{-11} kg m^{-2} s^{-1} and 6×10^{-11} kg m^{-2} s^{-1} respectively. After crushing the samples to expose fresh surface, the samples were leached at 25 °C for twenty consecutive weeks, yielding 10 ± 2 μg Ca per week, and a leach rate of $<1\times10^{-11}$ kg m^{-2} s^{-1}.

The reason for the difference in leach rates between the crystal and grains is not clear. It is not due to the tiny apatite inclusions (Table 1) in the perovskite crystal because our leach tests on apatite show similar leach rates to perovskite. We could find no evidence for extensive weathering or other impurities, such as calcite, in the crystal. We cannot, of course, absolutely rule out very minor, more soluble impurities affecting the leach rates. However, since glass leach rates decrease greatly with increasing SiO_2, it does not seem unreasonable that the low TiO_2 perovskite crystal (42.3% TiO_2) has a substantially higher leach rate than the high TiO_2 perovskite grains (58.01% TiO_2).

Table 2 summarises the leach data for the $SrTiO_3$ and $BaTiO_3$ synthetic samples. Several interesting points emerge. First, as for the above perovskite samples, the leach rate decreases markedly with leaching time. For $SrTiO_3$ at 100 °C, the leach rate (in kg m^{-2} s^{-1}) decreases from 1.5×10^{-8} for a one day leach to 1.9×10^{-9} for a two week leach (Table 2). Second, as for the two perovskite samples, the leach rates for both the Sr and Ba samples at all temperatures from 25 to 300 °C are within an order of magnitude of each other . Third, as for the perovskite grains, the leach rate at 100 °C is significantly higher than the leach rates at the lower and higher temperatures. Fourth, the leach rates on the synthetic perovskite analogues are similar to the crystalline $CaTiO_3$ results (Table 3), but considerably larger than those for the Western Australia perovskite grains.

The leach rates for perovskite and the Sr and Ba analogues are summarized in Table 3. Except for our Western Australia perovskite sample, the results were obtained on disk-like materials in static tests in similar volumes of water. Our results are at least qualitatively consistent with SYNROC leaching results[15] on powder samples[16]. The Ca leach rates reported earlier[16] (10 days leaching) are within an order of magnitude for all temperatures between 25 °C and 450 °C, and the leach rates drop markedly with time[16]. For example, as for our Sr sample at 100 °C (Table 2), the Ca leach rate from SYNROC at 100 °C drops by an order of magnitude from a 1 day to a 10 day leach. The leach rates reported[16] (8×10^{-12} to 6×10^{-11} kg m^{-2} day^{-1} from 25 °C to 450 °C) are certainly lower than our crystal and disk leach rates by as much as two orders of magnitude, but are rather close to our leach rates on perovskite grains. Powder

Table 2 Sr and Ba leach rates in perovskite analogues

Sample	Leaching T(°C)	Leaching time (days)	Surface area (cm^2)	1st week Sr or Ba (μg)	1st week Leach rate (kg m^{-2} s^{-1})	2nd week Sr or Ba (μg)	2nd week Leach rate (kg m^{-2} s^{-1})	3rd week Sr or Ba (μg)	3rd week Leach rate* (kg m^{-2} s^{-1})
$SrTiO_3$	25	7	2.1	300	2.3×10^{-9}	190	1.5×10^{-9}		
	100	1	2.1	265	1.5×10^{-8}				
	100	7	2.2	500	3.7×10^{-9}	170	1.3×10^{-9}	120	9.3×10^{-10}
	100	14	0.9	205	1.9×10^{-9}				
	200	7	2.4	100	7.0×10^{-10}	45	3.1×10^{-10}	ND†	$<6\times10^{-11}$
	300	7	2.0	60	5.0×10^{-10}	35	2.9×10^{-10}	ND	$<6\times10^{-11}$
$BaTiO_3$	25	7	2.1	140	1.0×10^{-9}				
	100	7	2.1	510	4.1×10^{-9}	450	4.1×10^{-9}	350	2.7×10^{-9}
	200	7	2.0	110	9.2×10^{-10}	110	9.2×10^{-10}	90	7.4×10^{-10}
	300	7	1.8	ND	$\leq4\times10^{-10}$	ND	$\leq4\times10^{-10}$	ND	$\leq4\times10^{-10}$

Samples were prepared by mixing the alkaline earth carbonate with TiO_2 for 3 h, calcining at 900 °C for 3 h, remixing for 3 h, calcining at 1,000 °C for 3 h, remixing, pressing and molding; before recalcining at 1,150 °C for 24 h. These samples weighed about 0.5 g, had geometric surface areas of ~2 cm^2, and had densities at least 90% of the theoretical values.

* Leach rates based on amounts of Sr or Ba in solution. Reproducibility is ±40%.

† ND, concentrations of Ba or Sr were below atomic absorption detection limits.

leach rates (using BET surface areas) can be as much as two orders of magnitude lower than disk leach rates using geometric surface areas[17,18].

Table 3 compares the leach rates for the titanates with those of rhyolite[19], and borosilicates[20]. Typical of a process controlled by kinetic factors, the glass leach rates increase dramatically with temperature. As a result, the leach rates for glasses at 200 °C or 300 °C are orders of magnitude larger than the titanates. However, at 100 °C and below the leach rates are similar: and at 25 °C it seems that, except for the Western Australia sample, the titanate leach rates are in fact larger than those for rhyolite[19] or higher melting (1,300 °C) borosilicates (K. B. Harvey, personal communication).

The unusual temperature dependence of the titanate leach rate is due to an interesting combination of the thermodynamics and kinetics. At 100 °C and 300 °C, the equilibrium Ca^{2+} concentrations for equation (13) at neutral pH are ~170 M (~7×10^6 μg per g or p.p.m.) and ~5×10^{-3} M(~200 p.p.m.) respectively. The relatively large leach rates at low temperatures are at least partially due to the large driving force provided by the extremely large equilibrium concentrations (equation (5.20) in ref. 8). The solubility of $Ca(OH)_2$, drops dramatically with increasing temperature from over 1,000 p.p.m. Ca^{2+} at 20 °C, to 370 p.p.m. at 100 °C to 20 p.p.m. (or 200 μg in 10 ml) at 250 °C (ref. 21). The solubility of the Sr and Ba hydroxides are likely to be similar. Indeed, $SrTiO_3$ and $BaTiO_3$ leached at 300 °C become coated with small crystals which we have observed by scanning electron microscopy. Preliminary X-ray diffraction, and electron spectroscopy for chemical analysis (ESCA) results provide evidence that these crystals are indeed the alkaline earth hydroxides which have reprecipitated on the surface.

Table 3 Leach rates of glasses and titanates (kg m^{-2} s^{-1})

	25 °C	100 °C	200 °C	300 °C
$CaTiO_3$ (crystal)	—	2.2×10^{-9}	—	8.9×10^{-9}
$CaTiO_3$ (grains)*	$<1\times10^{-11}$†	1×10^{-10}	—	1×10^{-11}†
$SrTiO_3$	2.3×10^{-9}	3.7×10^{-9}	7.0×10^{-10}	5.0×10^{-10}
$BaTiO_3$	1.1×10^{-9}	4.0×10^{-9}	9.2×10^{-10}	3.5×10^{-10}
Rhyolite‡	1.0×10^{-10}	4.1×10^{-9} 3.5×10^{-10}§	3.7×10^{-8}	1.5×10^{-7}
Na borosili‖ cates¶	$\geqslant6\times10^{-12}$ $10^{-9}-10^{-10}$	$\geqslant1.2\times10^{-9}$ 1.4×10^{-8}		

All results are given for the leaching of disks in the first week unless otherwise noted. The titanate leach rates are determined from Ca^{2+}, Sr^{2+} of Ba^{2+} in solution, while the glass results are determined from weight losses. The leach rates drop off drastically with time. For example, rhyolite and borosilicate leach rates can be in the 10^{-12} range after 3 month's leaching at 25 °C.

* These Western Australia grains had diameters of ~0.6 mm. The probe analysis gave: 37.66% CaO, 58.01% TiO_2 (total 95.67%) with small amounts of Na, Sr, Nb, Fe and other trace elements (see ref. 14).

† These leach rates will be comparatively low, because the sample had been leached before.

‡ Ref. 19.

§ Leach rates from Na (in solution not corrected for congruent dissolution).

‖ 30 day static leach (K. B. Harvey, personal communication). These glasses melted at ~1300 °C.

¶ Ref. 20.

Thermodynamics are clearly very important in controlling the leach rate. Moreover, it is not valid to do experiments at high temperatures and extrapolate to low temperature to obtain mineral leach rates using the kinetic model of glass leaching[1].

Perovskite geochemistry

Titanium is an accessory element of many common minerals[22] including sheet silicates (biotite), amphiboles (hornblende) and pyroxene (augite). It is an essential constituent of only a few minerals, such as sphene ($CaTiSiO_5$), ilmenite ($FeTiO_3$), rutile and its polymorphs (TiO_2) and perovskite ($CaTiO_3$). Ilmenite, sphene and rutile are not susceptible to alteration in sedimentary and hydrothermal environments and these resistate minerals may accumulate in sediments to be mined as mineral sands[23]. Despite the apparent stability of Ti-rich phases, one mineral, perovskite, may not be a resistate. Geological information supports this suggestion. Perovskite is a rare mineral formed only in rocks of unusual composition. In contrast, ilmenite, sphene, and rutile are common minerals found in rocks of extremely diverse composition, suggesting that they are stable relative to perovskite in most geological environments. Rocks containing perovskite are of plutonic and volcanic origin (ref. 22, p. 423), displaying exceptionally low SiO_2 and high CaO concentrations[5]. They have been generated in very high temperature–pressure environments, probably being produced in the mantle rather than in the Earth's crust (ref. 5, p. 526). We may therefore expect the rocks and constituent minerals to be unstable in upper crustal environments. Furthermore these rocks contain, in addition to perovskite, leucite ($KAlSi_2O_6$), kalsilite ($KAlSiO_4$) and nepheline ($NaAlSiO_4$) all of which are less stable than feldspars, in the presence of natural waters[10]. Their survival at the Earth's surface may result more from the volume of rock to be altered and from the arid climates in which they are found (Wyoming; Western Australia; southern Spain) than from their resistance to weathering and groundwater solutions. Perovskite is also observed in limestones thermally altered by granitic intrusions[22] but again the environment is unusual and very restricted. Minerals formed in such a setting may well be very unstable in other, more common, environments. From the geological evidence we conclude that perovskite will be stable only in unique geological settings. Our thermodynamic and kinetic studies strongly support this suggestion.

Conclusions

Contrary to earlier suggestions, we have found that perovskite is thermodynamically unstable in most rocks, and in natural waters. Perovskite appears to be kinetically far more stable than glasses at high temperatures ($T \gg 100$ °C). However, at the more important temperatures of $\leqslant 100$ °C, the comparison depends on the type of both glass and perovskite under consideration. For example, our best perovskite leach rate is lower than low melting borosilicates at all temperatures; but is comparable to, or greater than, the higher melting (1,300 °C) borosilicate at 25 °C. Our synthetic samples leach more rapidly at 25 °C than even some low melting borosilicates. Obviously, perovskite is not 'vastly' more stable kinetically than glass and silicates generally as has been claimed[1,2], and much more work is required to understand fully the kinetic stability and decomposition pathways of perovskite and other SYNROC minerals.

Ideally, minerals used as nuclear waste hosts should be thermodynamically stable. A mineral which is thermodynamically unstable and has a low laboratory leach rate, could well leach much faster in the field due to chemical or biological catalysts. Considerable efforts should now be made to examine the thermodynamics of mineral and glass solubilities considering the backfill and repository or vault materials.

We thank the Atomic Energy of Canada Ltd for financial support and Dr H. H. Schloessin and Dr B. Mason for perovskite samples. This paper was presented in part at the Workshop on Alternate Nuclear Waste Forms and Interaction in Geologic Media, held at Gatlinburg in May 1980.

Received 7 August; accepted 23 December 1980.

1. Ringwood, A. E., Kesson, S. E., Ware, N. G., Hibberson, W. & Major, A. *Geochem. J.* **13,** 141–165 (1979).
2. Ringwood, A. E. & Kesson, S. E., *Proc. int. Symp. Ceramics in Nuclear Waste Management*, CONF-790420 174–178. (US Department of Energy, Technical Information Center, Oakridge, 1979).
3. Mason, B. *Principles of Geochemistry*, 100–104 (Wiley, New York, 1966).
4. Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. U.S. geol. Surv.* **1452,** (1978).
5. Carmichael, I. S. E., Turner, F. J. & Verhoogen, J. *Igneous Petrology* (McGraw-Hill, New York, 1974).
6. Krauskaupf, K. B. *Introduction to Geochemistry* Ch. 4 (McGraw-Hill, New York, 1967).
7. Schuiling, R. D. & Vink, B. W. *Geochim. cosmochim. Acta* **31,** 2399–2411 (1967).
8. Fyfe, W. S., Price, N. J. & Thompson, A. B. *Fluids in the Earth's Crust* (Elsevier, Amsterdam, 1978).
9. Ellis, A. J. in *Geochemistry of Hydrothermal Ore Deposits* Ch. 11 (ed. Barnes, H. L.) (Holt, Rinehart and Winston, New York, 1967).
10. Helgeson, H. C. *Am. J. Sci.* **267,** 729–804 (1969).
11. White, D. E., Hem, J. D. & Waring, G. A. *Data of Geochemistry* 6th edn, Ch. F (Geological Survey Professional Paper 440-F, 1963).
12. Nesbitt, H. W. & Bricker, O. P. *Geochim. Cosmochim. Acta* **43,** 403–409 (1978).
13. Erickson, R. L. & Balde, L. V. *Geol. surv. Profess. Pap.* No. 425 (1963).
14. Mason, B. *N. Z. Dept Sci. Ind. Res. Bull.* **218,** 114–120 (1977).
15. Reeve, K. D. *et al. Proc. Waste Management '80 Symp.* (University of Arizona, 1980).
16. Tewhey, J. D. *Workshop on Alternate Waste Forms and Interactions in Geologic Media*, Gatlinburg (1980).
17. Hench, L. L. & Clark, D. E. *J. Non-Crystal. Solids* **28,** 83–105 (1978).
18. Ross, W. A., Turcotte, R. P., Mendel, J. E. & Rusin, J. M. *Proc. int. Symp. Ceramics in Nuclear Waste Management* CONF-790420, 52–56 (US Department of Energy, Technical Information Center, Oakridge 1979).
19. Karkhanis, S. N., Bancroft, G. M., Fyfe, W. S. & Brown, J. D. *Nature* **284,** 435–437 (1980).
20. Boult, K. A., Dalton, J. T., Hall, A. R., Hough, A. & Marples, J. A. C. *Proc. int. Symp. Ceramics in Nuclear Waste Management* CONF-790420, 248–255 (US Department of Energy, Technical Information Center, Oakridge, 1979).
21. Linke, W. F. *Solubilities of Inorganic and Metal Organic Compounds* (American Chemical Society, 1965).
22. Deer, W. A., Howie, R. A., & Zussman, J. *An Introduction to the Rock Forming Minerals* (Longmans, London, 1966).
23. Hutton, C. O. *Bull. geol. Soc. Am.* **61,** 635 (1950).

The Norse merman as an optical phenomenon

W. H. Lehn
Department of Electrical Engineering, University of Manitoba, Winnipeg, Canada R3T 2N2
I. Schroeder
60 Thatcher Drive, Winnipeg, Canada R3T 2L3

Mediaeval Norse writings that describe the merman are considered accurate observations of a natural phenomenon. Images of common sea mammals, severely distorted by strong, non-uniform atmospheric refraction, fit the mediaeval descriptions remarkably well. Three examples are presented: computer-generated images of a killer whale and a walrus, and a photograph of a suitably distended boulder. The mediaeval correlation of a merman sighting with the advent of storms is also verified.

MEDIAEVAL manuscripts contain occasional passages, apparently deriving from invention or superstition, for which a basis can be found in observation or experience. However, some texts, such as those that describe the merman, have defied positive identification. In the seventeenth century it was first suggested that the merman was simply a dugong or manatee[1], a view that was generally accepted by the end of the nineteenth century[2]. Although this interpretation is very unsatisfactory, it has remained the most prevalent theory[3].

In general, modern scientists have avoided the study of such anomaly sightings, partially because it is regarded as a non-soluble problem, and partially because it can be so easily distorted[4].

This article considers the merman, as conceived by twelfth and thirteenth century authors. The natural source of the merman legend will be identified, and the high level of sophistication attained by the Norse in observing and correlating natural phenomena will be demonstrated. Finally, the concept of the merman will be traced through its subsequent evolution.

Mediaeval accounts

The *King's Mirror*, written in the mid-thirteenth century, is noted for its highly accurate descriptions of sea mammals and other natural phenomena[5]. The section dealing with the North Atlantic contains only three items that assume an aspect of the marvellous: the *hafgerdingar* and the merman, found in the Greenland Sea, and the *hafgufa*, or kraken, in the Iceland Sea[6]. The explanation of the *hafgerdingar* (sea fences) as a visual effect created by anomalous atmospheric refraction[7] suggests that the merman be reconsidered.

The author of the *King's Mirror* describes the merman as:

> "This monster is tall and of great size and rises straight out of the water.... It has shoulders like a man's but no hands. Its body apparently grows narrower from the shoulders down, so that the lower down it has been observed, the more slender it has seemed to be. But no one has ever seen how the lower end is shaped, ... No one has ever observed it closely enough to determine whether its body has scales like a fish or skin like a man. Whenever the monster has shown itself, men have always been sure that a storm would follow."

He proceeds to discuss its mate, the mermaid, which differs from the merman in that it possesses heavy hair, breasts, large webbed hands, and a fish's tail, but also "rarely appears except before violent storms".

An earlier account of the merman is found in the *Historia Norvegiae*, written[8] about 1170:

0028–0836/81/050362—05$01.00

"there is the whale, and there is the merman, the largest wild beast, but without head and tail, thus it is just like a treetrunk when it leaps up and down, and it does not show itself without its foretelling danger for seafarers."

These two passages, the only descriptions of the merman that have survived from mediaeval times, will be analysed here. First the merman's shape will be considered as an optically distorted image; then the correlation between such a sighting and the advent of storms will be briefly assessed.

Atmospheric optics

The atmosphere occasionally produces optical distortion strong enough to make ordinary objects unrecognizable, even at fairly close distances. Such distortions have been suggested as the source for many of the sightings, both ancient and modern, of lake and sea monsters[9]. Common sea mammals, so distorted, fit the mediaeval merman descriptions.

The underlying optical theory, here briefly summarized, is given elsewhere[10–12]. Light rays travelling nearly horizontally follow paths whose vertical curvature κ depends on the distribution of atmospheric density ρ with elevation z. The curvature is given by

$$\kappa = \frac{\beta}{1+\varepsilon\rho}\,\frac{d\rho}{dz}$$

where $1+\varepsilon\rho$ is the refractive index of air ($\varepsilon = 0.000226$). The density is related to the more easily measured absolute temperature T by[13]

$$\rho = \frac{\beta p_0}{T(z)}\exp\left[-g\beta\int_0^z \frac{dz'}{T(z')}\right]$$

where p_0 is the surface pressure, g the acceleration due to gravity, and $\beta = 3.49\times10^{-3}$, a constant of proportionality. Temperature inversions, with their associated stable atmosphere, cause a downward refraction of light rays. Irregularities in the temperature profile, especially thermoclines, create the irregularities in refraction necessary for severe image distortion.

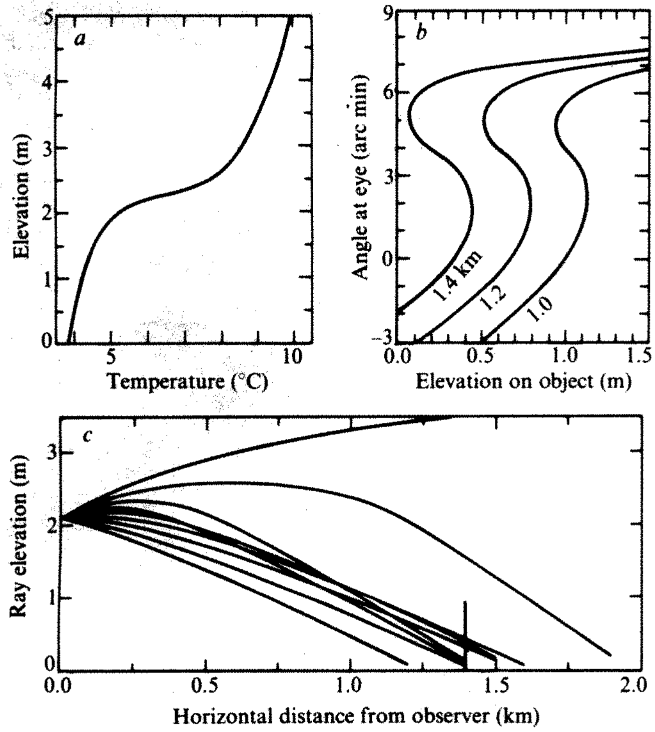


Fig. 1 *a*, Temperature profile used in *Example 1*. *b*, Transfer characteristics generated by the profile of *a*, for object distances of 1.0, 1.2 and 1.4 km. Eye elevation is 2.1 m. *c*, Paths described by light rays in temperature conditions of *a*. The uppermost ray enters the eye at an angle of +9′; the lowest with −3′. The angular increment between successive rays is 1.5′; the spacing of 0.5′ used to plot *b* is too dense to produce a meaningful plot at this scale. The ordinate measures ray elevation above the (curved) surface of the sea. The location and height of the whale's head (see Fig. 2*a*) is indicated by the vertical bar at 1.4 km.

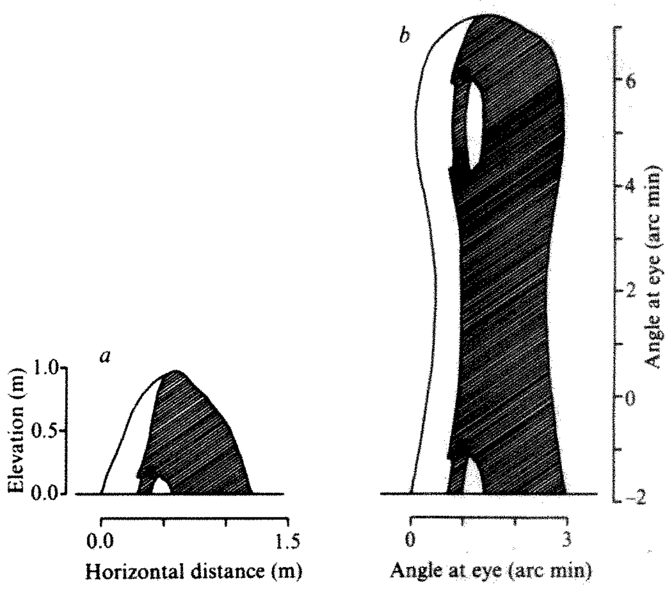


Fig. 2 *a*, A killer whale, spy-hopping. *b*, The distorted image of the whale, seen from a distance of 1.4 km in the temperature conditions of Fig. 1*a*. If the same temperature profile were to prevail for some distance beyond the whale, then the *hafgerdingar* would be present; the apparent horizon would be above the top of the image, at an elevation of +10′, about 10 km from the observer. For a normal atmosphere these values are −2.5′ and 5.7 km.

Computer programs developed for ray tracing[13,14] have been used to identify temperature profiles that create merman images. The reconstructions that follow are based on inversions with thermoclines situated a few metres above the sea.

Example 1: A killer whale projecting its head vertically out of the water is an excellent source for merman images. Such activity, called 'spy-hopping', is common among cetaceans of many sorts[15,16]. It is also specifically mentioned by Shackleton[17] as a hunting procedure used by killer whales[18].

In the image-construction process, several quantities can be manipulated: shape of temperature profile, elevation of observer's eye and distance from observer to object. After some computer experimentation with all of these elements, the following typical conditions were chosen to illustrate the effect. Figure 1*a* shows the temperature profile: a moderate inversion of total strength 7.5 °C and with a thermocline at elevation 2.2 m. The atmosphere is assumed laterally homogeneous over the short ranges involved, so that this temperature profile prevails everywhere between object and observer. If the observer's eye is now placed at an elevation of 2.1 m, the light rays entering the eye follow the paths shown in Fig. 1*c*. The elevations at which these rays intersect an object plane at some distance from the eye are plotted as a transfer characteristic in Fig. 1*b*. These curves determine the nature of the image perceived by the observer. The eye assumes each incoming ray to be straight, and projects it back on to the object plane to give an (apparent) image point. A complete set of curved rays, as in Fig. 1*b*, thus causes a redistribution of all the object points on the object plane. As the ray bending is in a purely vertical plane no lateral image distortion is present.

The original and distorted images are compared in Fig. 2*a* and *b*, respectively, for an object distance of 1.4 km. Figure 2*a* is adapted from a photograph of a killer whale[19]. The distorted image in Fig. 2*b* is unlikely to be recognized as a whale.

Space restrictions mean that the many possible variations of the image cannot be displayed. With changes in the determining parameters, the shape can appear more or less slender; the midsection can become quite a narrow 'neck' or 'waist'; and the white eye-spot may not be imaged in the 'head'.

Example 2: A similar image-constructing process was carried out for a walrus. As this is a smaller animal, a shorter range was

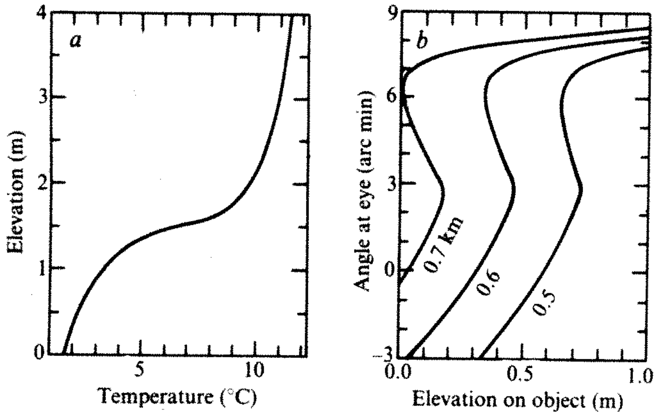


Fig. 3 *a*, Temperature profile used in *Example 2*. *b*, Corresponding transfer characteristics, for object distances of 0.5, 0.6 and 0.7 km. Eye elevation is 1.45 m.

chosen (0.7 km) to have a sufficient horizontal angle subtended at the eye. A steeper temperature profile was required to produce the necessary refraction over the reduced range. Figure 3*a* shows the temperature profile, an inversion of strength 11 °C with a thermocline at about 1.5 m elevation. The transfer characteristics corresponding to an eye elevation of 1.45 m are shown in Fig. 3*b*.

Figure 4 compares the original with the distorted image. A walrus photographed in its natural habitat[20] forms the basis for Fig. 4*a*. The distorted image for the stated conditions is in Fig. 4*b*. Again many variations are possible in this distortion: the figure can become stockier or narrower; it need not have a thin 'waist'; and the apparently fierce 'mouth' and 'teeth' at the top of the figure can become much larger.

Example 3: A photograph of a present-day merman is given in Fig. 5*b*. It appeared over Lake Winnipeg on 2 May 1980, a hot calm day, while there was still some ice on the lake. The inland temperature at the time was about 28 °C, while the lake surface was near 0 °C. The source of the apparition, subsequently identified, was an insignificant boulder (Fig. 3*a*) sited 1.10 km from the camera.

Comparison with mediaeval descriptions

The author of the *King's Mirror* describes the merman with great care, while the less well-informed author of the *Historia Norvegiae* provides only a very general description. Yet the overall impression is identical.

Both accounts describe the merman as a huge animal rising straight out of the water. The lack of detail in the *Historia Norvegiae* implies that the animal had only been seen at some distance, a fact confirmed by direct statement in the *King's Mirror*. The enormous size could, therefore, only have been estimated by comparing its appearance with that of known objects at apparently similar distances.

The man-like shape implied by its name is consistent with the visual impression of an object viewed at some distance, when vertically elongated. As the horizontal dimension remains unchanged, the resulting image acquires a large height-to-width ratio, a form identified with man.

The accounts diverge only in the detailed configuration and in the specific features. The *Historia Norvegiae* suggests a cylindrical body, while the *King's Mirror* gives it shoulders like a man's, with a tapering body. As shown in Figs 2*b*, 4*b* and 5*b*, both shapes are a natural consequence of refractive distortion. Further, the *Historia Norvegiae* claims the merman has no head at all, while the *King's Mirror* lists a pointed head, having eyes, mouth, nose, chin and neck. Again, the process of refraction permits various upper terminations, and refracted markings in approximately the right place (Figs 2*b*, 4*b*) can create the impression of a face. As small atmospheric irregularities tend to disintegrate the image, producing a whiskery or hairy effect[9], such characteristics could be attributed to its mate, the mermaid. At the distances implied in the *King's Mirror*, the human eye, whose resolving power is between 0.5 and 1 arc min, registers only the major outlines; finer details are interpolated by the perception process. Eventually, numerous sightings similar in overall impression but differing somewhat in detail would result in a generalized description and the identification of a species.

The elevation of the observer is important in all optically induced anomaly sightings. The refractive distortion generally diminishes if the image can be viewed from above the thermocline. However, such variations would go unnoticed by Norse mariners, for the working of their ships was carried out from the deck, a few metres above the sea[21]. The subsequent use of much higher decked ships and the establishment of elevated lookouts would explain the infrequent merman sightings in later centuries.

Finally, the distorted boulder of Fig. 5*b* bears a striking resemblence to the *umibohzu*, or "priest of the sea", a giant-like monster known to the Japanese[22]. An even stronger likeness occurs in a narrow-necked version of Fig. 2*b*, in which the *umibohzu*'s eyes are produced by the whale's white eye patch. Clearly the Japanese were interpreting the same visual phenomena in much the same way as did the Norse.

Correlation with the environment

The thermoclines that generate merman images are best created when a warm air mass slowly moves over significantly cooler surface air. In conditions of relative calm, the boundary between the two air masses will possess the flatness and uniformity necessary for well-defined (but distorted) images.

A typical situation is illustrated by the last stages of a warm front, when the warm–cool interface has almost descended to the surface. This has been experimentally verified by Wegener's observations of good superior mirages in the North Atlantic[11,23], which he later correlated with the arrival of warm fronts[24]. However, neither he nor subsequent polar researchers have studied the fine structure of the temperature distributions that the lowest air layers must possess to produce merman images.

Scoresby[25], an experienced Arctic whaler, recorded numerous examples of anomalous refraction which strongly resemble the marvels of the *King's Mirror*: the elevated horizon, a dense appearance of the atmosphere, and the vertically elongated hulls of ships as if in a heeling position. Though Scoresby recognized some forms of refraction phenomena and correlated them with an easterly wind, he attributed others to vapours in the atmosphere, or confused them with ice-blink. He thus associated unusual visual effects with storms but did not identify

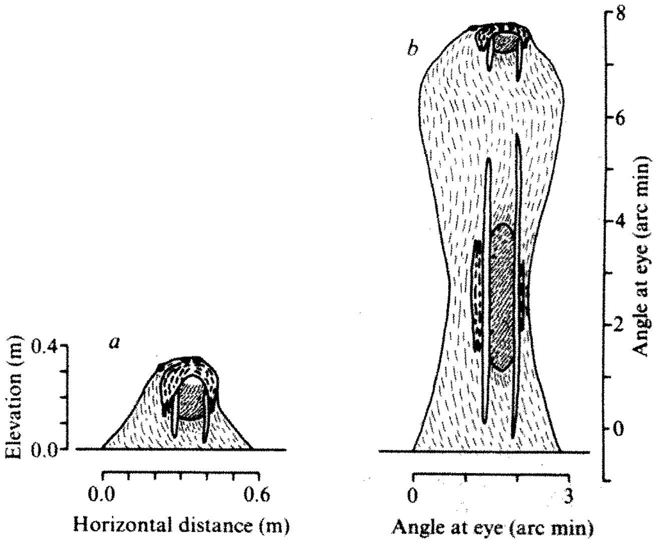


Fig. 4 *a*, Walrus. *b*, The distorted image of the walrus, seen from a distance of 0.7 km in the temperature conditions of Fig. 3*a*. Again, the *hafgerdingar* will be present if the same temperature profile exists over a sufficiently wide area. The horizon elevation would be +13′, at a distance of 15 km, compared with −2′ at 4.8 km for a normal atmosphere.

them with refraction. The dead calm followed by a sudden rise in temperature, the typical conditions noted by Scoresby just before a major storm, are ideally suited to the development of a thermocline. I have observed such an effect on a very localized scale, in the Canadian Arctic, over the Beaufort Sea: the sudden appearance of a medium-range mirage, followed within 40 min by a sudden temperature rise, rainstorms, and disappearance of the mirage. The amount of optical distortion depends directly on the temperature difference between the two air masses, which in turn determines the strength of the front and the severity of the subsequent storms.

Similar refractive distortions will have provided the Norse with sufficient opportunity to observe elongated images of the sea mammals which they used as navigational guides[26]. Because Severin's experience in the *Brendan* indicates that whales do not avoid small sailing vessels[27], the limiting factor would have been the frequency of suitable atmospheric conditions rather than the presence of a suitable subject. These observations belonged to the thirteenth century mariner's general body of knowledge, and were recorded as such in the *King's Mirror*. It is fitting that the *King's Mirror* places both phenomena, the *hafgerdingar* and the merman, together in its text. From the optical reconstruction, as well as Scoresby's observations, clearly some form of *hafgerdingar* will be observed along with the merman. Both are correctly associated with the advent of storms on the open sea, and are located in the Greenland Sea where such extreme conditions were also observed by Scoresby.

Subsequent development

The precise descriptions in the *King's Mirror* directly contrast with those of Olaus Magnus three centuries later. His monsters, composites of several species assembled from fragmented reports[19], reflect the general level to which Norse maritime knowledge had declined. As knowledge of the North Atlantic was regained by western European explorations, unrecognizable sightings reminiscent of the merman were occasionally reported[28,29]. Nevertheless, the Norse merman had become largely a legendary creature, and skepticism as to its existence hardened when scientists identified it with the newly discovered dugong and manatee, both warm water mammals. To clarify this problem, Pontoppidan collected written accounts, as well as verbal reports from reliable witnesses. His attempt to isolate the factual from the mythical established the North European merman as an existing species with some well-defined characteristics. This study formed the basis for all future work even though Pontoppidan did not have access to the *King's Mirror*, a manuscript long believed lost[1].

It is of interest to examine the evolution of the thirteenth century merman into Pontoppidan's eighteenth century composite. Not only had the natural habitat of the merman shifted from the Greenland Sea to the coastal waters of northern Europe, but the conditions in which the merman was sighted had also changed. Many of the sightings accepted by Pontoppidan were reported on very warm, calm days. These observations suggest that the characteristic vertically distended appearance of an optically distorted image was still being identified as a merman; for such weather conditions, though different from those of the advancing front, also develop strong temperature gradients and consequent refraction. Further, while the specific correlation with frontal activity on the open sea had totally disappeared, the fear of the merman had survived as a generalized bad omen. Finally, the species had become approachable; whereas the early Norse reported only distant sightings, actual catches and findings of partially decomposed remains were claimed in the seventeenth and eighteenth centuries.

A similar evolution can be traced for the *hafgufa*. As originally described, it was a featureless monster of enormous dimensions in the Iceland Sea[6]. So few sightings had ever been reported that the author of the *King's Mirror* estimated its world population at one or perhaps two, and hesitated to include it in a compendium of sea life. By the eighteenth century, this monster had multiplied remarkably to inhabit the entire North Atlantic, especially the coastal waters of Norway. Pontoppidan recorded the varied accounts of this creature but found the information so diverse and inconsistent that he did not attempt to identify it with any known species[1]. However, by the end of the nineteenth century, the kraken had assumed the characteristics of a giant squid, an idea which has become increasingly accepted[2].

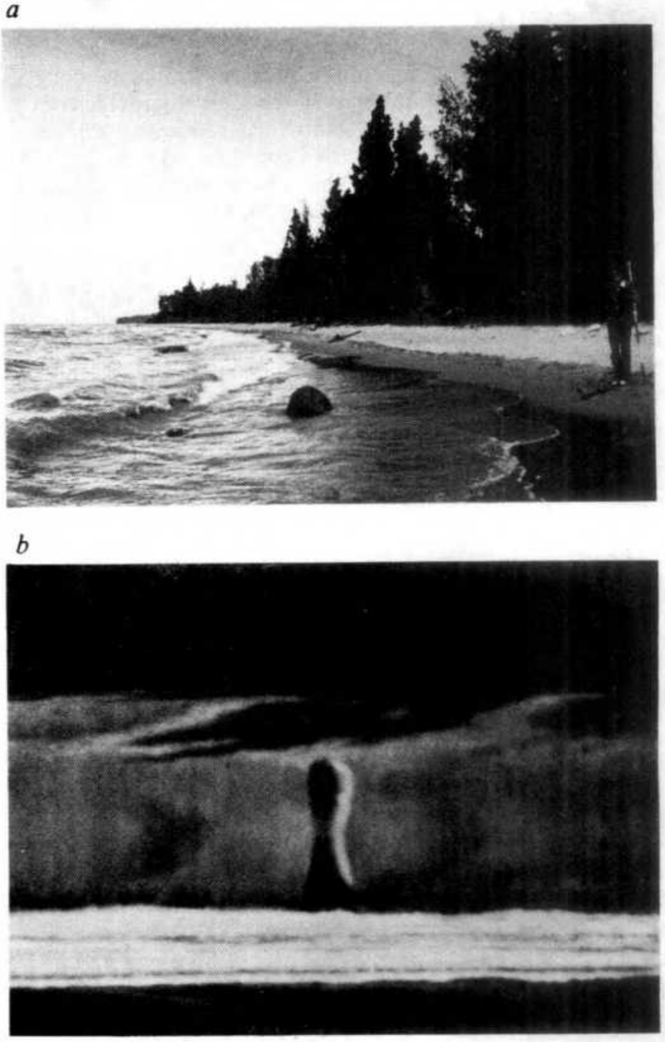


Fig. 5 *a*, Lake Winnipeg, 28 May 1980. The boulder in centre foreground was identified as the source of the image in Fig. 5*b*. It is ~68 cm wide and should show 30–35 cm above water on a calm day. *b*, The boulder distorted into a merman shape, 3:17 p.m. CDT (central daylight time), 2 May 1980. The image subtends 5.4′ vertically and 2.2′ horizontally. The photograph was made from a distance of 1.10 km using a mirror lens (focal length 1,250 mm) on Kodachrome 64 film. the exposure was 1/250 s at *f*/11, and the lens elevation was 2.5 m above the lake. A photograph taken just previously, at 3:15.5 p.m., shows the distortion as incompletely developed. The merman image of this boulder lasted for only a few minutes. Such conditions occur almost every spring on Lake Winnipeg.

The *King's Mirror* is so accurate in its description of the merman and the *hafgerdingar* that a serious effort should be made to identify a natural phenomenon closely fitting the original *hafgufa*. Its appearance suggests that it may be a consequence of submarine volcanic activity in Icelandic waters. Steenstrup[26] used this theory to explain the *hafgerdingar*. But the *hafgerdingar* are much better described as an optical phenomenon, while Steenstrup's theory is a superior match, both in appearance and location, to the thirteenth century *hafgufa*.

Following the deterioration of contact between Iceland, Greenland and Norway, the merman and the *hafgufa* had become culturally isolated. Cut off from their origins, these stories were free to develop according to local conditions. Progress in identifying these marvels was hampered by the fact that they had evolved in time, a problem similar to the one identified in a recent analysis of the Gunnbjorn Skerries[30]. When renewed interest in the North Atlantic stimulated scholars to seek out all previously known information, both the skerries and the marvels of the *King's Mirror* emerged as composite remnants from the mediaeval age.

Conclusion

The merman described in the early sources, like the *hafgerdingar*, is, in fact, an accurate description of an unusual optical phenomenon. Its appearance is exactly that predicted by computer simulations of anomalous atmospheres, and the weather conditions with which it is associated agree with modern observations in arctic regions. The correspondence of these two marvels with reality leads directly to the conclusion that the *hafgufa* account must also be considered as an observed natural phenomenon.

The extent to which Norse expertise deteriorated, once trans-Atlantic navigation ceased, is clearly demonstrated by the many unrealistic accounts found in sixteenth century sources. Clear and precise reports were superseded by confusion and the supernatural. The tendency of modern scholars has, therefore, been to dismiss as inaccurate or superstitious any observations inconsistent with their own experience. The optical interpretation of the merman and the *hafgerdingar* enables a clear differentiation between thirteenth and sixteenth century concepts and furthers an appreciation of the visual impressions which formed an integral part of Norse maritime experience.

This research was supported in part by the University of Manitoba's Research Board and Northern Studies Committee, as well as by NSERC grant A3196. We thank Dr F. W. Hougen for checking our Norwegian translation.

Received 21 July; accepted 14 November 1980.

1. Pontoppidan, E. *The Natural History of Norway*, II, 186–195, vi, 210–218 (Lind, London, 1755).
2. Oudemans, A. C. *The Great Sea Serpent* (Brill, Leiden, 1892).
3. Heuvelmans, B. *In the Wake of Sea Serpents* (Hill and Wang, New York, 1968).
4. Westrum, R. in *The Sociological Review Monograph No. 27* (ed. Wallis, R.) 293–314 (University of Keele, 1979).
5. Nansen, F. *In Northern Mists*, II, 243 (Heinemann, London, 1911).
6. Larson, L. M. (transl.) *The King's Mirror*, 135–138, 125 (Twayne, New York, 1917).
7. Lehn, W. H., Sawatzky, H. L. & Schroeder, I. *Scand. Rev.* **66**, 2, 36–41 (1978).
8. Koht, H. (transl.) *Historia Norvegiae* (*Den eldste Noregs-Historia*), 13 (det Norske Samlaget, Oslo, 1921).
9. Lehn, W. H. *Science* **205**, 183–185 (1979).
10. Pernter, J. M. & Exner, F. *Meteorologische Optik* 2nd edn (Braumüller, Vienna, 1922).
11. Koch, J. P. & Wegener, A. *Meddr Grønland* **75**, 609–629 (1930).
12. Lehn, W. H. *J. opt. Soc. Am.* **69**, 776–781 (1979).
13. Lehn, W. H. & Sawatzky, H. L. *Polarforschung* **45**, 120–128 (1975).
14. Lehn, W. H. & El-Arini, M. B. *Appl. Opt.* **17**, 3146–3151 (1978).
15. Nickerson, R. *Brother Whale*, 101–102, 117 (Chronicle, San Francisco, 1977).
16. Slijper, E. J. *Whales and Dolphins*, 119 (University of Michigan Press, Ann Arbor, 1976).
17. Shackleton, E. H. *South—The Story of Shackleton's Last Expedition 1914–1917*, 22, 35, 323 (Macmillan, New York, 1920).
18. Norman, J. R. & Fraser, F. C. *Field Book of Giant Fishes*, 289–294 (Putnam, New York, 1949).
19. Matthews, L. H. *The Whale*, 248, 22–32 (Allen & Unwin, London, 1968).
20. Bruemmer, F. *Encounters with Arctic Animals*, 73 (McGraw-Hill Ryerson, Toronto, 1972).
21. Falk, H. S. *Altnordisches Seewesen*, 67 (Winter, Heidelberg, 1912).
22. Sweeney, J. B. *A Pictorial History of Sea Monsters*, 49 (Crown, New York, 1972).
23. Wegener, A. *Meddr Grønland* **42**, 125–356 (1914).
24. Wegener, A. *Annaln. Hydrogr. Berl.* (Köppenheft), 93–95 (1926).
25. Scoresby, W. *An Account of the Arctic Regions, with a History and Description of the Northern Whale-Fishery*, I, 387–391, 442, 300, 398–399 (1820; reprinted by Kelley, New York, 1969).
26. Jones, G. *The Norse Atlantic Saga*, 115, 146 (Oxford University Press, London, 1964).
27. Severin, T. *The Brendan Voyage*, 124–132, 156, 188 (Arrow, London, 1979).
28. Best, G. *The Three Voyages of Martin Frobisher* (ed. Stefansson, V.) I, 100–101, lvi (Argonaut, London, 1938).
29. Gould, R. T. *The Case for the Sea Serpent*, 12–16 (Philip Allan, London, 1930).
30. Lehn, W. H. & Schroeder, I. *Polarforschung* **49**, 173–187 (1979).

Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution

I. A. Wilson, J. J. Skehel* & D. C. Wiley

Gibbs Laboratory, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138
*National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The haemagglutinin glycoprotein of influenza virus is a trimer comprising two structurally distinct regions: a triple-stranded coiled-coil of α-helices extends 76 Å from the membrane and a globular region of antiparallel β-sheet, which contains the receptor binding site and the variable antigenic determinants, is positioned on top of this stem. Each subunit has an unusual loop-like topology, starting at the membrane, extending 135 Å distally and folding back to enter the membrane.

INFLUENZA virus acquires a lipid membrane during maturation by 'budding' through the host-cell membrane. The resulting viral membrane contains the two virally coded integral membrane glycoproteins, haemagglutinin and neuraminidase[1]. The haemagglutinin (HA) of the A/Hong Kong/1968 virus is a trimer of 224,640 molecular weight (MW). It may be solubilized from the viral membrane by bromelain digestion, which removes a 5,406 MW C-terminal hydrophobic (anchoring) peptide from each subunit[2–4]. This observation, extended by the results of subsequent primary sequencing experiments[5–14], places the haemagglutinin in a class of integral membrane proteins characterized by a three-domain structure with a large hydrophilic, carbohydrate-containing domain on the external surface of the membrane, a small, uncharged hydrophobic peptide of 24–28 amino acids spanning the membrane, and a small, hydrophilic domain (10–55 amino acids) on the internal side of the membrane (Fig. 1*a*). Other well characterized examples of this class of membrane glycoprotein include the human and murine histocompatibility antigens HLA and H-2 (refs 15, 16), erythrocyte glycophorin[17], and the animal virus membrane proteins. Such proteins show enzyme or 'recognition' activities, in contrast to proteins active in membrane transport, of which the structurally best defined example is bacterial rhodopsin[18,19].

Typical of membrane glycoproteins, the haemagglutinin chain is initially synthesized to include an N-terminal hydrophobic 'signal' peptide[11,20,21], which is subsequently removed as part of the process by which the protein is transported across and anchored into the membrane[22,23] (Fig. 1*a*). Each polypeptide chain of the trimer is glycosylated at seven sites with a total carbohydrate of 13,000 MW (19% by weight)[6,11]. We wish to understand the structural correlates of these properties. What is the nature of the membrane attachment of this protein? What is the location and biological role of the carbohydrate? As a virus membrane protein, the haemagglutinin expresses a number of surface activities whose structural bases are also sought. How

0028-0836/81/050366—08$01.00

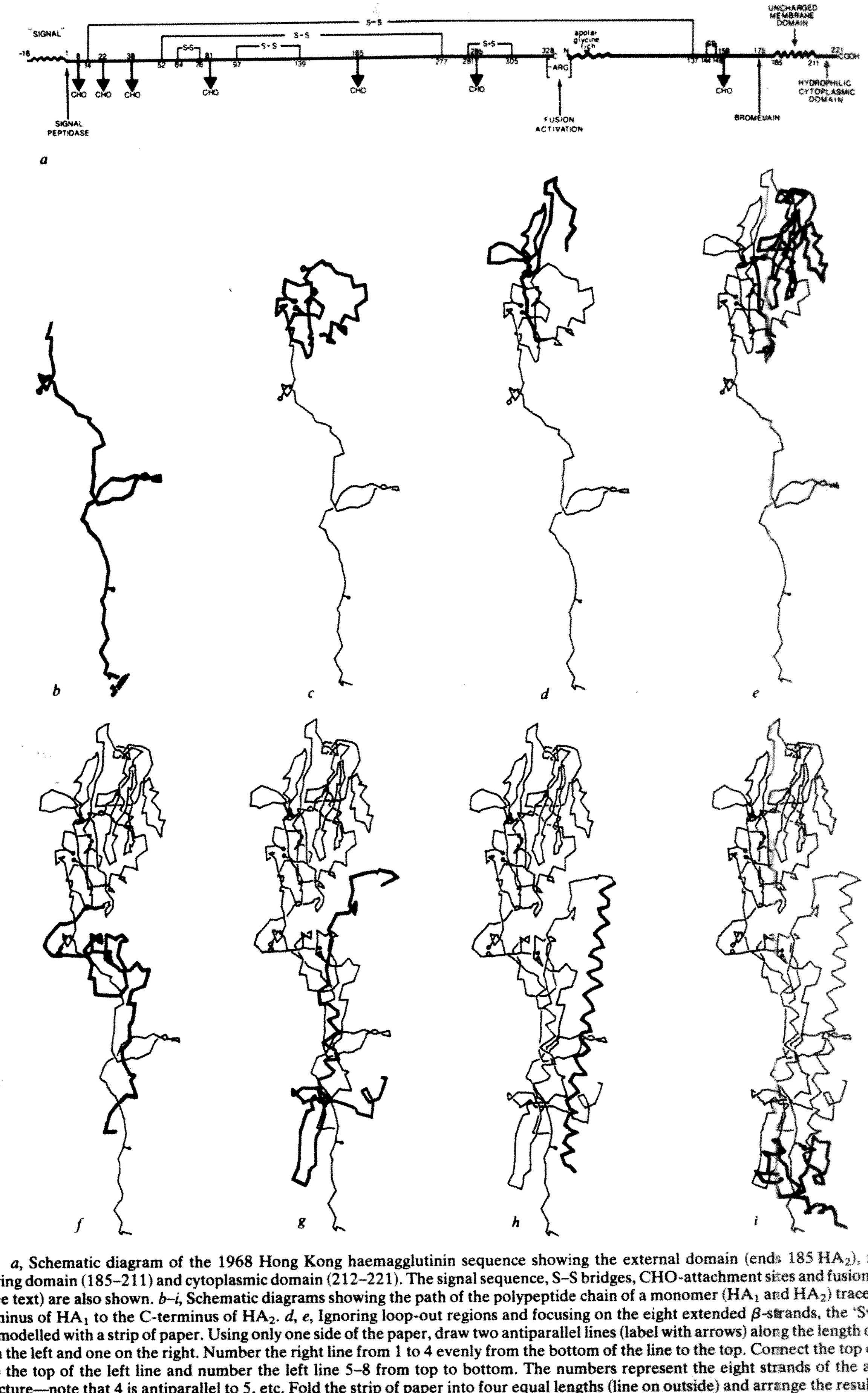


Fig. 1 *a*, Schematic diagram of the 1968 Hong Kong haemagglutinin sequence showing the external domain (ends 185 HA_2), membrane anchoring domain (185–211) and cytoplasmic domain (212–221). The signal sequence, S–S bridges, CHO-attachment sites and fusion activation site (see text) are also shown. *b–i*, Schematic diagrams showing the path of the polypeptide chain of a monomer (HA_1 and HA_2) traced from the N-terminus of HA_1 to the C-terminus of HA_2. *d, e*, Ignoring loop-out regions and focusing on the eight extended β-strands, the 'Swiss roll' is easily modelled with a strip of paper. Using only one side of the paper, draw two antiparallel lines (label with arrows) along the length of the strip, one on the left and one on the right. Number the right line from 1 to 4 evenly from the bottom of the line to the top. Connect the top of the right line to the top of the left line and number the left line 5–8 from top to bottom. The numbers represent the eight strands of the antiparallel β-structure—note that 4 is antiparallel to 5, etc. Fold the strip of paper into four equal lengths (line on outside) and arrange the resulting 'Swiss roll' so that strands 1, 8, 3 and 6 are visible in that order from one side and 2, 7, 4 and 5 are seen from the other (Fig. 2). The point between strands 4 and 5 corresponds to amino acid 208 (HA_1) in the haemagglutinin structure (Fig. 3*a*). The eight strands are approximately 1, 119–123; 8, 251–259; 3, 174–182; 6, 229–237; 2, 163–170; 7, 241–248; 4, 200–205 and 5, 210–213 (Fig. 3*a*). (This figure is intended as a description of the structure, not a folding pathway.)

does this protein recognize the sialic acid-containing host-cell receptor[24]? How does the post-translational removal of an arginine (329), dividing the chain into two disulphide-linked chains, HA_1 (328 amino acids) and HA_2 (221 amino acids)[11,14], activate infectivity and membrane fusion activity[25–27]? How does variation in the amino acid sequence of the structure alter the antigenic determinants of the molecule and lead to recurrent epidemics of respiratory disease?

We report here a first description of the results of the 3Å-structure determination of the external, glycoprotein domain of the influenza virus haemagglutinin, which provides initial answers to certain of the above questions. The host-receptor binding site and antigenic determinants are located on a globular region which lies on top of a long fibrous coiled-coil (Fig. 2). The fusion activation peptide is located near (35 Å) the virus membrane end of the molecule. The polypeptide fold is unusual, like a long loop, starting at the membrane, extending 135 Å distally, and ending at the membrane. This tertiary structure indicates that the molecule is probably extruded from the membrane as a single chain during biosynthesis and may even remain attached via the signal peptide to the membrane during folding. Most of the oligosaccharide chains are found near the membrane end of the molecule. The identification of antibody binding sites and the contribution of their variation to the recurrence of respiratory disease in man is discussed in the accompanying article[28].

Structure determination

The bromelain-released haemagglutinin (30 mg ml^{-1}) crystallizes in tetragonal space group $P4_1$, $a = 163.2$ Å, $c = 177.4$ Å from 1.28 M sodium citrate, 0.1% azide, *p*H 7.5 (ref. 29). A single isomorphous derivative, mercury phenylglyoxal[30], was used in structure determination. 1° oscillation photographs (12-h exposure) were recorded at 4 °C, using an Elliot G-X-6 rotating anode source operated with a 100 μm focus at 40 kV and 20 mA, and Franks focusing optics[31]. Up to 15° of data could be recorded from each crystal by exposing a fresh volume to X rays after 24 h exposure (Table 1).

Table 1 Data collection statistics

Data set	No. films	Total reflections measured	Unique reflections	R_{sym}	R_{whole}	R_{asym}
Native	78	408,672	83,015	0.12	0.11	0.18
Derivative	62	338,738	78,077	0.13	0.13	0.20

$R = \sum |s_i I_{hi} - \bar{I}| / \sum \bar{I}_n$ where s_i = scale factor for crystal *i*, I_{hi} = intensity of reflection *h* of crystal *i* and $\bar{I}_n$ = average intensity of reflections. R_{sym}, symmetry-related reflections on the same film; R_{whole}, fully recorded reflections from different films; R_{asym}, all reflections (≥50% recorded) on different films, final scaling.

Solution of the 3 Å difference Patterson revealed six heavy-atom locations per molecule (Table 2*a*), two sites 27 Å apart per monomer, which define a non-crystallographic, molecular 3-fold axis. (Details of the structure determination will be published elsewhere.) The asymmetric unit contains one glycoprotein trimer[32] of 208,422 MW and is 78% solvent (by weight). The standard isomorphous replacement (SIR) phases of 59,669 reflections were determined to 3 Å resolution with an overall figure of merit of 0.34 and overall phasing power of 1.38 (Table 2*b*). Eleven cycles of refinement of these and previously undetermined phases of a total of 74,194 reflections by non-crystallographic 3-fold symmetry averaging[33] dramatically improved the quality of the electron density maps and the agreement of calculated and observed amplitudes ($R = 0.47$ cycle 1, $R = 0.27$ cycle 11). The molecular envelope required redetermination three times during refinement to include protruding regions of protein and carbohydrate that had been truncated by the original estimate of the envelope but that were clearly apparent in difference Fourier maps.

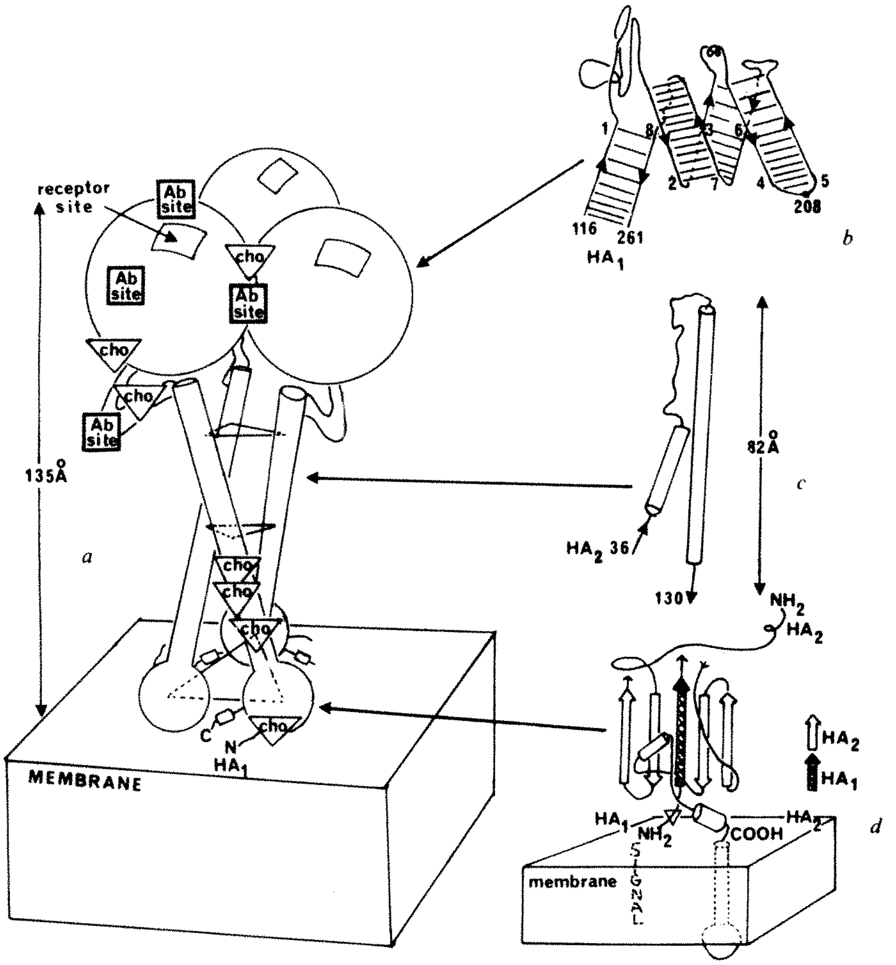


Fig. 2 *a*, Schematic diagram of the 1968 haemagglutinin trimer showing carbohydrate attachment sites (CHO), antigenic sites (Ab site) (see ref. 28), the host–receptor binding site and the arrangement relative to the membrane. *b*, The eight-stranded β-sheet structure and looped-out region in the globular domain. *c*, HA_2 residues 36–130, including two α-helices (cylinders), form part of the fibrous region. Three of the long helices, one from each monomer, pack together as a triple-stranded coiled-coil which stabilizes the trimer. *d*, The membrane end of the molecule contains a five-stranded β-sheet. The central strand is the N-terminal of HA_1 and the adjacent strands come from the C-terminal chain of HA_2. Broken lines suggest the path of the hydrophobic anchoring peptide cleaved by bromelain. The possible attachment of the N-terminus to the membrane by the signal sequence is also indicated. The site of the first oligosaccharide chain (residue 8) is shown as a triangle.

Interpretation of electron density map

A minimap (4 Å cm^{-1}) was calculated for the trimer, sectioned normal to the 3-fold axis. The subunit boundaries were clearly visible. Ambiguities in chain tracing were resolved with the aid of the amino acid sequence (provided by Fiers *et al.*[11] and Ward *et al.*[14]). α-Carbon positions of residues 4–328 of HA_1 and residues 1–175 of HA_2 have been located. The amino terminus of HA_1 and the C-terminus of HA_2 are close to the membrane end of the haemagglutinin and therefore to the bromelain cleavage site at residue 175 of HA_2 (ref. 34). The unaveraged map indicates that deviations from 3-fold symmetry may occur in the amino-terminal region of HA_1 by interaction of one of the amino-termini with an adjacent molecule in the crystal. The side-chain densities are well defined and unambiguous for 50–60% of the residues and another 20–30% fit moderately well. Only 15% of the residues have weak side-chain or main-chain density, and the chain can clearly be traced throughout the entire molecule.

Location of the six disulphides showed that only one disulphide linkage connects HA_1 and HA_2 (refs 7, 35).

Description of the structure

The haemagglutinin trimer is an elongated cylinder, 135 Å long, with a triangular cross-section varying in radius from 15 to 40 Å. The structure is divided into two distinct regions: (1) a long fibrous region, containing residues from both HA_2 and HA_1, is centred on a 76 Å triple-stranded α-helical coiled-coil extending from the membrane surface; (2) a globular region containing residues entirely from HA_1 sits on top of the fibrous stem and includes an 8-stranded β-structure.

Assembly of the trimer seems to be a requirement for stabilization of the elements of the subunit tertiary structure in the fibrous stem region.

Tertiary fold. The unusually extended tertiary structure of the haemagglutinin may be appreciated by following the schematic drawing of its chain tracing (Fig. 1). The first 63 amino acids reach from the N-terminus at the membrane end of the molecule in an almost extended conformation 96 Å along the molecular length before becoming involved in a compact fold (Fig. 1*b*, *c*). The fold (Fig. 1*c*) contains a short disulphide loop (64–76) and a three-turn helix.

β-Structure. The distal end of the molecule forms an 8-stranded antiparallel β-sheet structure from residue 116 to 261 (Fig. 1*d*, *e*; see also Fig. 3). Similar 'Swiss roll' β-structures have been described in tomato bushy stunt virus[36] and southern bean mosaic virus[37], and have been discussed by Richardson[38]. Two looped-out regions between strands in the β-structure seem to be important antigenic regions: the 125–163 loop between β-strands 1 and 2, and the 187–199 α-helix-containing loop between β-strands 3 and 4 (ref. 28). This β-structure also provides the framework for the sialic acid-receptor binding site (see below).

Switch region. The globular β-region reconnects to the fibrous stem by only two disulphide loops 52–277 and 281–305 (Fig. 1*f*). The globular region is thus an HA_1 loop connected to the fibrous stem by only two antiparallel chains, the N-terminal chain extending away from the membrane and the C-terminal chain returning towards the membrane. In the fibrous region the C-terminal chain of HA_1 extends 60 Å towards the membrane end of the molecule, remaining antiparallel to the amino-terminal HA_1 chain.

HA_1–HA_2 junction. The N-terminus of HA_2 is 21 Å from the C-terminus of HA_1, indicating a substantial rearrangement in this region when the haemagglutinin is activated by cleavage (Fig. 1*g*). The HA_2 amino-terminal glycine-rich sequence forms an uncommon series of four contiguous reverse turns before wrapping transversely around the fibrous coil region ~35 Å from the membrane end of the molecule.

Fibrous helices. Two antiparallel α-helices form the backbone of the fibrous region of each subunit. A short α-helix (29 Å) extends distally from the membrane and connects by an extended chain to a very long (53 amino acid, 14 turn) α-helix which stretches 76 Å back towards the membrane (Fig. 1*g*, *h*). In the trimer, this long helix twists around two identical helices from other subunits to form the coiled-coil core of the fibrous region. The conformation of the extended chain (HA_2, 59–76) and the first half of the long α-helix (HA_2, 77–100) seems to be stabilized predominantly by extensive contacts with the other subunits of the trimer (Fig. 3).

Terminal globular region. The membrane end of the fibrous region contains a compact globular fold which includes a 5-stranded antiparallel β-sheet. The structure of this sheet is noteworthy—the central strand (no. 3) is the amino-terminal chain of HA_1, and the extended chains of HA_2, 350 amino acids away in the sequence, form the two strands flanking it on each side (nos 1, 2 and 4, 5) (Figs 1*g*–*i* and 2). The chain terminates in a three-turn α-helix at the bromelain cleavage site, Gly 175 (ref. 34), at the extreme membrane end of the molecule (Fig. 1*i*). An uncleaved haemagglutinin would enter the membrane 10 amino acids further along. At this stage in the analysis we cannot rule out the possibility that the C-terminal helix was normal to the membrane and may have rearranged as a result of bromelain digestion.

The complete HA_1 chain contains about 36% β-sheet, 6% α-helix and 16% reverse turns. The HA_2 chain has a significantly more regular secondary structure with 14% β-sheet, 53% α-helix and 15% reverse turns.

Biosynthesis and folding. The preceding description of the tertiary structure is not intended to suggest a pathway for the folding of the protein. The acquisition of the three-dimensional fold of a molecule that has been synthesized in the cytoplasm and that has had most of its mass transported to the external side of the membrane is poorly understood.

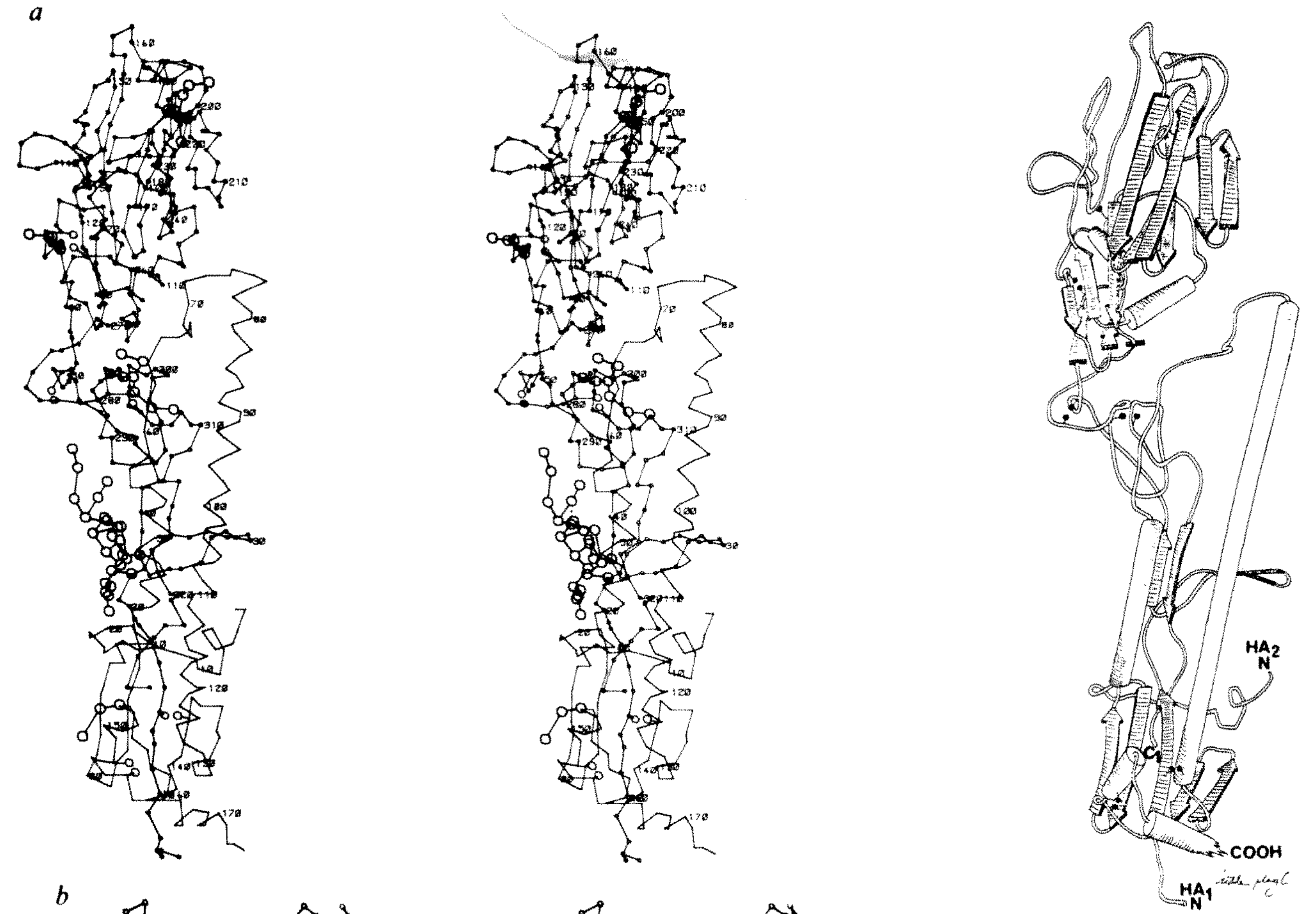
The tertiary fold of the haemagglutinin structure suggests that the molecule was extruded from the membrane as a single chain

Table 2 Solution of 3 Å difference Patterson with SIR phase refined parameters and statistics

a	Atom site	x	y	z (Å)	Occupancy	B (Å^2)
	1	30.9	53.4	0.0*	0.47	18.6
	2	25.1	54.0	21.9	0.60	13.4
	3	0.4	43.0	−16.5	0.36	16.4
	4	3.8	57.7	38.6	0.38	20.8
	5	13.0	41.6	7.3	0.47	19.2
	6	41.7	78.5	1.0	0.35	17.4

b	Resolution (Å)	8.7	6.9	5.7	4.8	4.2	3.7	3.3	3.0	Total
	No. reflections	989	2,223	3,968	5,726	8,206	11,431	13,859	13,267	59,669
	$\bar{M}$	0.40	0.41	0.38	0.36	0.36	0.34	0.33	0.31	0.34
	F_c/E_i	1.54	1.74	1.69	1.58	1.47	1.39	1.27	1.10	1.38
	R_{KRAUT}	0.11	0.13	0.14	0.13	0.13	0.15	0.18	0.24	0.16

Occupancy, relative fractional occupancies (not absolute). *B*, isotropic temperature factor.
* Atom 1 z fixed at 0.0.



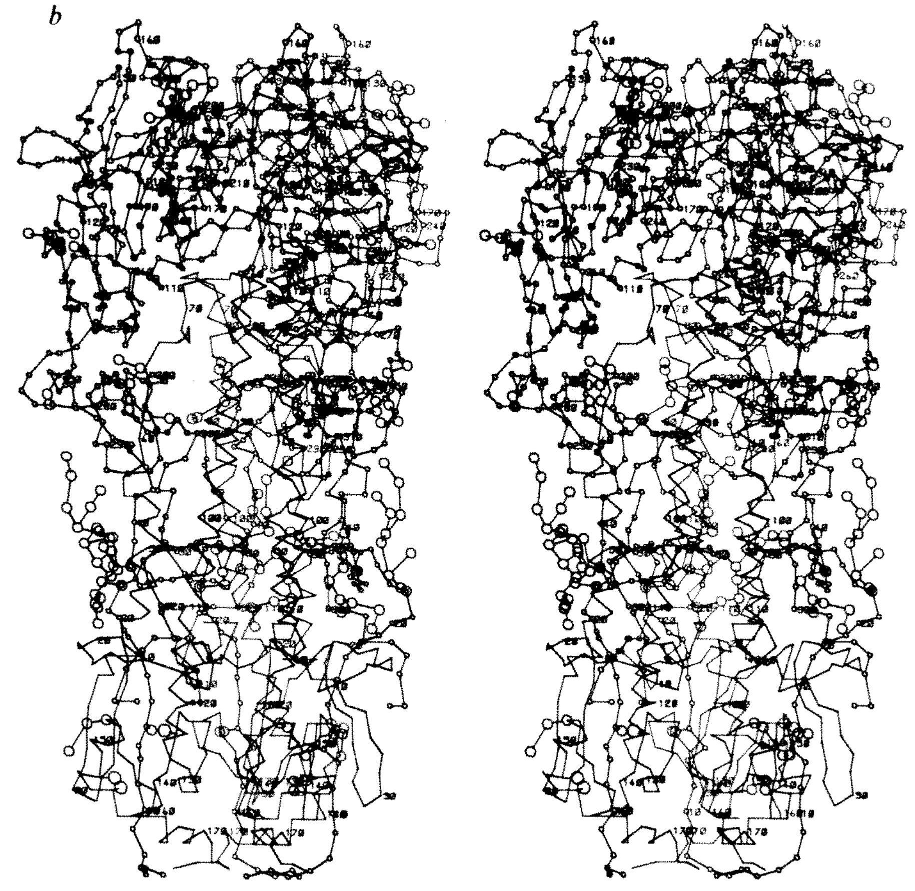


Fig. 3 *a*, Stereo drawing of the α-carbon tracing of an influenza haemagglutinin monomer: ∘, HA_1; –, HA_2; hexose, ○. We expect the present estimate of hexose ring positions to change when stereochemically precise model building and refinement is completed (see text for further explanation). The schematic diagram on the right shows flat, twisted arrows for extended β-chains, cylinders for helices, and filled circles for disulphide bonds. *b*, Stereo drawing of the haemagglutinin trimer: ∘, HA_1; –, HA_2; hexose, ○. The molecule is 135 Å long and 15–40 Å in radius.

(a globular region found at the end of a 'structureless' first 63 amino acids) as opposed to a more cooperative refolding that can be imagined if the entire molecule emerged from the membrane by a concerted conformational change[23,39].

Furthermore, the loop-like topology, with both the N- and C-termini of the chain at the extreme membrane end of the molecule, suggests that the structure may have folded, attached to the membrane not only by the carboxy-terminal anchor peptide (or extruding chain) but also by the amino-terminal hydrophobic signal sequence. (In this respect, the shortest uncharged section in all known haemagglutinin signal sequences (A/Memphis/10/78[H1N1])[40] is 11 amino acids long, sufficient

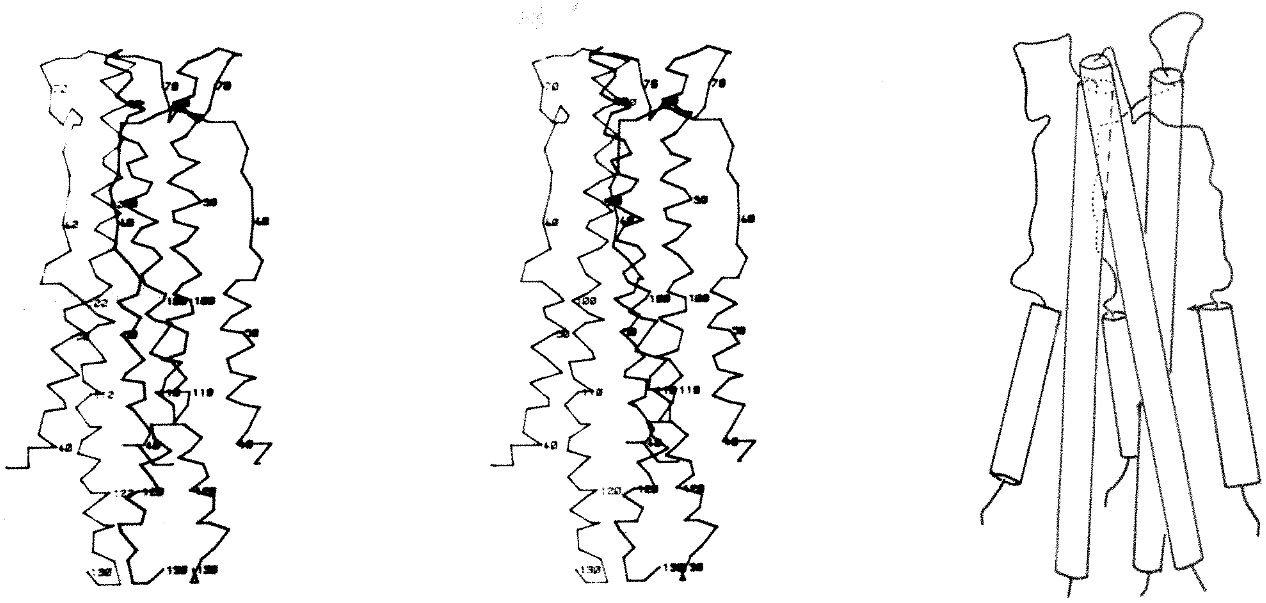


Fig. 4 Stereo drawing of amino acids 36–130 of the HA_2 chain, illustrating the triple-stranded coiled-coil core of the trimer. The schematic drawing on the right illustrates the same region with cylinders representing helices.

to span the lipid bilayer in an extended conformation.) Although the peculiar tertiary structure of the haemagglutinin could be generated without either of these suggestions, they emerge naturally from a consideration of the molecular form.

Quaternary structure

Coiled-coil. The major forces stabilizing the haemagglutinin's trimeric subunit interactions arise from a triple-stranded coiled-coil[41,42] in the fibrous region of the molecule. An α-helix (74–126) extends 76 Å along the molecular 3-fold symmetry axis. Three such helices, one from each subunit, pack together and twist 100° around each other in a left-handed superhelix (Fig. 4). The N-terminal (top) half of the coiled-coil is tightly packed (10 Å helix to helix) with many nonpolar residues in van der Waals contact around the 3-fold axis (Ile 77, Leu 80, Val 84, Leu 91, Leu 98, 99, 102). The C-terminal half of the coiled-coil is expanded away from the 3-fold axis (up to 22 Å helix to helix), with polar and charged residues interacting around the 3-fold axis (His 106, Asp 109, Asp 112, Ser 113, Asn 116, Glu 120, Arg 123). Some polar interactions may be mediated by solvent molecules.

Numerous buried charge pairs form intra-chain and inter-subunit salt bridges that contribute to the stability of the coil. The amino-terminal residues Leu and Phe of HA_2 pack around the 3-fold symmetry axis in the beginning of the open region of the coil (112–118).

The coiled-coil interactions agree with theoretical predictions of helix packing residues based on a heptad repeat[43], but the a, d contact region drifts through a, e, to b, e, by the end of the helices. Ward and Dopheide[44] predicted a coiled-coil in the approximately correct location from heptad repeats in their haemagglutinin sequence data.

Three short α-helices (HA_2, 38–56) pack almost hexagonally and antiparallel with the lower half of the three long helices (Fig. 4). The interactions again involve a twisting such that one end of the short helix lies between two long helices and the other end lies against the long helix of its own subunit.

Trimer assembly. Assembly of the trimeric 'spike' provides stabilizing interactions which seem to be crucial for the tertiary folding of regions in the fibrous stem of the individual subunits. The top half of the long α-helix (HA_2, 77–100) has very few interactions within its own subunit and in a monomer would expose a 36 Å hydrophobic exterior surface. Similarly, the extended chain, HA_2 65–77, and the β-loop of HA_1, amino acid 30, make most of their stabilizing contacts with an adjacent subunit. Thus the polypeptide chain folding in these regions of the monomer may only be completed on assembly of monomers to form the trimers. A corollary of these observations is that the molecule may only project off the surface of the membrane as a 'spike' when assembled as a trimer. Whether the possible flexibility in the monomer stem region in the absence of trimeric contacts may be exploited (by assembly–disassembly) for any biological role remains to be tested.

Membrane-associated region. The top half of the triple-stranded coiled-coil in the fibrous region of the haemagglutinin suggests a model for the quaternary interactions that may take place in the uncharged, predominantly hydrophobic, C-terminal peptide of the complete haemagglutinin (Fig. 1*a*), the domain that anchors the molecule into the virus membrane. This membrane-associated peptide (185–211) is probably α-helical. Twenty-seven uncharged amino acids are sufficient to cross the lipid bilayer in a helical conformation and the hydrogen bonding along the helix would prevent the unfavourable occurrence of unpaired carbonyl oxygen and amide nitrogen atoms in the lipid environment. Whether the three chains of the haemagglutinin trimer cross the membrane together as a unit or separately is unknown. A triple-stranded coiled-coil of α-helix crossing the membrane as a unit is suggested as it would facilitate trimerization on the internal side of the membrane, a property which may be important in the mechanism of budding or other transmembrane functions.

Globular region trimer contacts. The quaternary interactions of the globular region are far fewer than in the fibrous region. Residues from the eight-stranded β-structure are in contact but the interface is not extensive (Fig. 3*b*). The oligosaccharide attached at residue 165 (HA_1) spans the interface between the subunits, forming substantial contacts on both sides that would stabilize the oligomer in this region.

Carbohydrate

The carbohydrate (MW 13,000 per monomer) is located on the surface of the trimer, all along its length, with more in the half of the molecule nearest the membrane surface than at the distal end (Fig. 3). Seven Asn-X-Thr/Ser (six in HA_1) sequences occur in the Hong Kong haemagglutinin and all are glycosylated[6,11]. Two of the oligosaccharide moieties are simple and contain only *N*-acetyl glucosamine and mannose (HA_1, 165,285) whereas the others also contain galactose and fucose (HA_1, 8,22,38, and HA_2, 154). The absence of terminal sialic acids is due to the activity of the viral neuraminidase. The carbohydrate on residues HA_1 165 and 285 is clearly visible in the electron density maps, whereas residues 22 and 38 (HA_1) have less well defined carbohydrate density and residues HA_1 8, 81 and HA_2 154 have

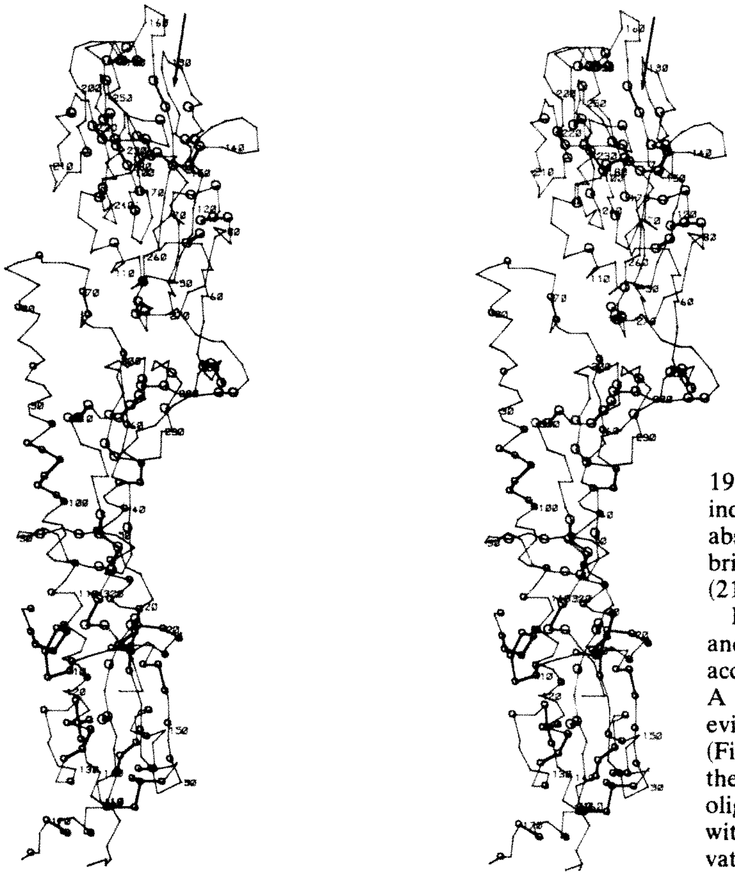


Fig. 5 Conserved amino acid positions from the known haemagglutinin sequences: ○, HA_1; ∘, HA_2. The pocket of highly conserved amino acids at the distal (top) end of the molecule, tentatively identified as the host-cell oligosaccharide receptor binding site, is marked by an arrow.

little or no visible carbohydrate in present maps, possibly due to positional disorder.

We have assigned preliminary positions to 63% of the hexose rings of the oligosaccharides. Our analysis is at present incomplete; for example, our placement of hexose rings is inconsistent with expected branch points at some locations (Fig. 3). Based on the experiences of Huber and colleagues with an oligosaccharide chain on IgG, we expect that when stereochemically precise model building and refinement are completed, these structures will be reliably determined[45,46].

The carbohydrate visible in the map covers 17–20% of the available surface and external residues of the trimer. Residues covered by the carbohydrate are mainly hydrophilic (60%), with Ser and Thr (33%) most frequently interacting with the sugars. Hydrophobic amino acids are often covered (33%) but no clusters of hydrophobic amino acids occur around any substituted Asn residue.

The role of carbohydrate in membrane antigens is largely undefined. Oligosaccharide chains on the haemagglutinin seem to have some structural roles, an observation also made for an IgG oligosaccharide[45]. Carbohydrate–protein interactions are involved in the subunit–subunit interface and inter- and intrachain contact regions. Thus these interactions may stabilize the monomer and the trimeric structure. We cannot assess the possibility that contacts between trimers on the viral surface could involve oligosaccharide chains. Two lysines, three arginines and five aromatic residues, potential tryptic and chymotryptic sites, are covered by oligosaccharide chains which suggests that carbohydrates may also stabilize the haemagglutinin by providing protection against degradative enzymes. The possibility that carbohydrate modulates recognition of the haemagglutinin by the immune system is discussed in the accompanying paper[28].

Structure comparison

Other influenza haemagglutinins seem to have structures very similar to that which we have found for the 1968 Hong Kong strain. The haemagglutinins from 12 strains of influenza have been sequenced (some partially) including viruses from 1957 to 1978 and an avian influenza[5–14,47]. Sequence comparison indicates that 22% of HA_1 and 45% of HA_2 residues are absolutely conserved—these residues include the six disulphide bridges and many structurally important residues such as proline (21 residues) and glycine (9 residues).

Furthermore, the structure reveals that all known insertions and deletions in the sequences of these different strains can be accommodated at the surface of the 1968 Hong Kong structure. A core of structurally important conserved residues is also evident in both the globular and fibrous regions of the molecule (Fig. 5). Similarly, all the carbohydrate attachment sites found in the known sequences are on the surface and could accommodate oligosaccharide side chains, and detailed structures associated with antigenicity also seem to be conserved[28]. These observations indicate that the 1968 structure is a good first-order model for haemagglutinins from many strains. Deviations from this structure will be important, however, in understanding the antigenic variation which the virus undergoes.

Biological activities

The distribution of conserved amino acids on the 3 Å structure reveals a highly conserved region in a surface pocket on the distal end of the molecule that seems well suited for a binding to host-cell oligosaccharide (Fig. 5). This tentative receptor binding pocket includes conserved residues Tyr 98, His 193, Glu 190, Trp 153 and Leu 194 (Fig. 5). It is similar to the wheat-germ agglutinin sialic acid-binding site[48]. No direct experiment has associated virus–cell receptor binding with specific amino acids in the haemagglutinin. Crystallographic data on the haemagglutinin in the presence of oligosaccharides are being collected to locate the receptor binding site directly.

The most highly conserved sequence in the haemagglutinin is the amino terminus of HA_2, where 1 substitution in the first 11 and only 5 in the first 23 amino acids have been observed. This sequence is associated with the activity by which the virus penetrates a host-cell membrane to initiate infection. Cleavage into HA_1 and HA_2 at this point is required for infectious penetration and for *in vitro* membrane fusion[26,27]. An homologous sequence is present at the cleavage-activation site on the Sendai virus fusion protein, where cleavage is required both for infectivity and membrane fusion activity[49]. The HA_2 N-terminal sequence is strongly nonpolar, the first charged side chain occurring at position 11 (Glu). It is unexpectedly rich in glycines (1,4,8 (ref. 12) and 13,16,20,23 (refs 4, 13, 14)) which could indicate flexibility or the presence of an unusual conformation. In the present structure, the first 11 amino acids form a contiguous series of sharp bends, with residues 2 and 3 buried, making 3-fold contacts at the centre of the molecule (Figs 3*b*, 5). (The series of glycine-rich bends in our unrefined model have not been identified with any known form of helix.) The α-amino group forms a salt bridge to an aspartic acid. The strong conservation of the apolar sequence is particularly striking when compared with signal or C-terminal anchoring hydrophobic sequences. In both the latter cases, almost every position has accepted conservation apolar substitutions

St. Groth and colleagues[15,16,23] has involved the selection of variants by growth of an H3 1968 virus (A/NT/60/68) and its derivatives in the presence of the most avid fractions of hyperimmune antisera and, similarly, Laver, Webster and colleagues[22,24,25] have selected variants of A/Memphis/1/71, a virus closely related to A/Hong Kong/1/68, by growth in the presence of monoclonal antibodies prepared against its haemagglutinin. After virus selection, amino acid sequence data were obtained for the haemagglutinins of the variants. Although only a partial sequence data are available, they indicate that variants selected using monoclonal antibodies contain single amino acid substitutions, which directly or indirectly alter an antigenic region of the haemagglutinin molecule sufficiently to prevent neutralization of the variant by the selecting antibody[22]. The locations of the amino acid substitutions in haemagglutinins of the laboratory-selected variants are summarized in Fig. 2*c*.

Suggested locations for the antigenic sites. An antigenic site is a region of a molecule involved in antibody binding[26]. Although there are no exhaustive immunochemical data such as those used in studies defining the antigenic sites on lysozyme and myoglobin[27,28], a combination of the amino acid sequence information and the three-dimensional structure of the 1968 haemagglutinin[1] allows identification of four regions of the Hong Kong haemagglutinin molecule in which amino acid substitutions seem to affect antibody binding.

Site A: an unusual protruding loop from amino acids 140 to 146, (see Fig. 2) which projects 8 Å from the local molecular surface, forms the centre of the most obvious antibody-binding site. Clearly the haemagglutinin of each antigenically distinct virus of epidemic significance has a mutation in this region (Fig. 2 and Table 1): A/Memphis/102/72: residue 144, Gly to Asp, and residue 122, Thr to Asn[4,6]; A/Victoria/1/75 and A/Victoria/3/75: residue 145, Ser to Asn, and residue 137, Asn to Ser respectively[9,12,13]. Furthermore, the haemagglutinins of a group of monoclonal antibody-selected variants neutralization by single amino acid substitutions in this region: 143, Pro to His, Ser, Thr, Leu; 144, Gly to Asp, and 133, Asn to Lys[22]. Conserved amino acids Gly 134, Ala 138, Cys 139, Phe 147 and Phe (Tyr) 148 occurring at each end of the antigenically variable loop provide the structural foundation of the site in all 10 haemagglutinins examined (including one Avian[3] and one Asian[7] subtype).

Site B: this comprises the external residues 187–196 of an *α* helix, and adjacent residues along the upper edge of a pocket tentatively implicated in virus receptor binding[1] (see Fig. 2). Two substitutions in the haemagglutinin of the 1972 virus, A/Memphis/102/72: 188, Asn to Asp, and 155, Thr to Tyr[4,6]; and two more in that of the 1975 virus, A/Victoria/3/75: 189, Gln to Lys and 193, Ser to Asn[9,12,13] characterize the site, but the lack of sequence information for this region of the haemagglutinins of several variants leaves the comparison incomplete[9]. The substitution 186, Ser to Ile in the haemagglutinin of the 30D virus[16], selected using 'avid-fraction' antisera, may also support the site designation. In addition, the existence of this structure in other haemagglutinins is suggested by the observed conservation of neighbouring residues. On the inward face of the helix residues 190 Glu, 191 Gln, 194 Leu and 195 Tyr are conserved in all the known sequences, as are 153 Trp and 154 Leu. However, because no amino acid substitutions in the haemagglutinins of variants selected by monoclonal antibodies comparable with those which support the delineation of site A have yet been reported, more direct evidence of antibody binding to this region is not available.

Site C: a bulge in the tertiary structure at the disulphide bond between Cys 52 and Cys 277, 60 Å from the distal tip of the molecule, comprises another antibody-binding site (see Fig. 2). The haemagglutinins of the viruses from both epidemic periods have substitutions clustered in this region:

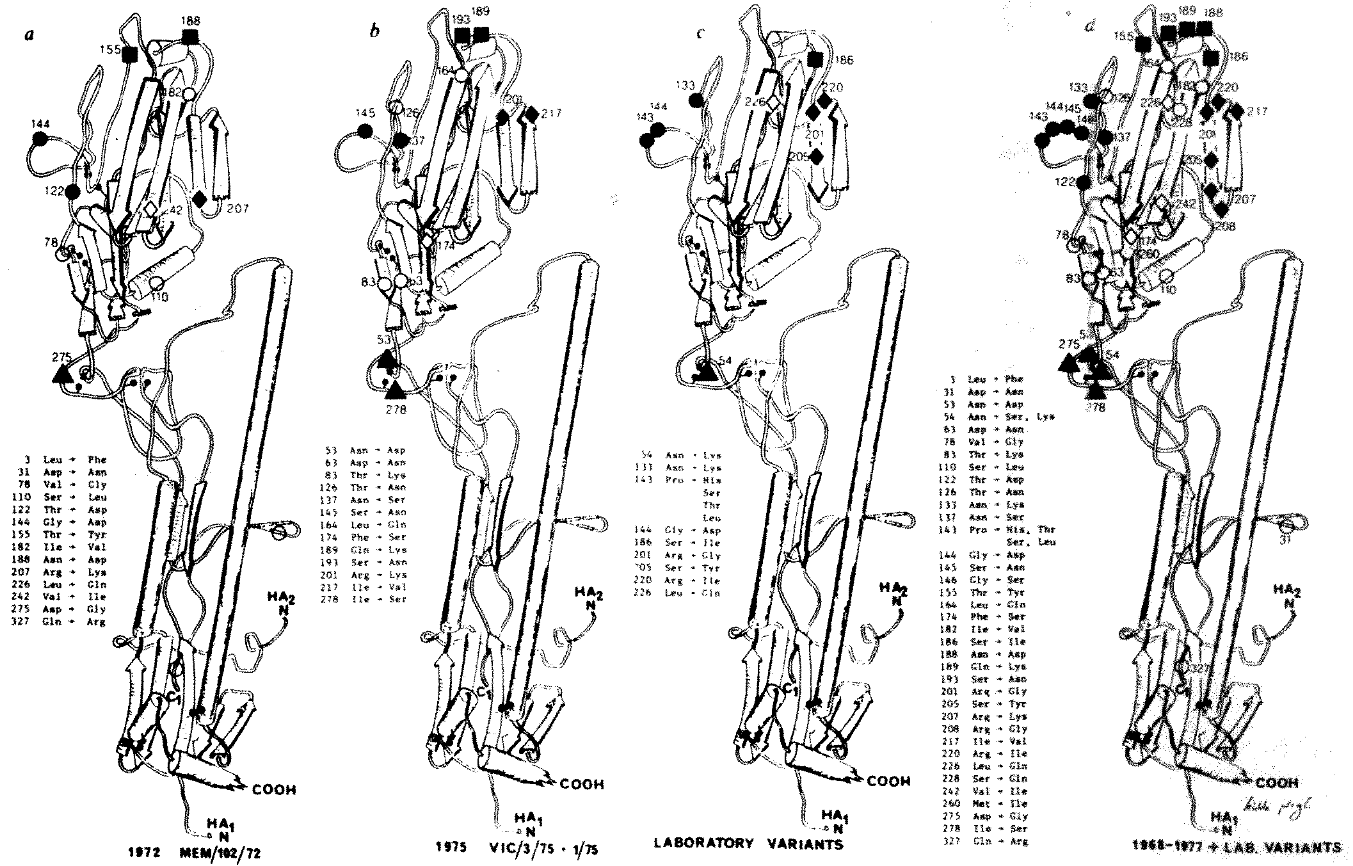


Fig. 2 Schematic drawing of a monomer of the AICHI-1968 haemagglutinin showing locations of amino acid substitutions on HA_1. ●, Site A; ■, site B; ▲, site C; ◆, site D (subunit interface); ◇, surface; ○, neutral. *a*, Amino acid changes from AICHI/2/68 to MEM/102/72 (refs 4–6, 10, 11). *b*, Amino acid changes from MEM/102/72 to VIC/3/75 (refs 12, 13). *c*, Laboratory-selected antigenic variants from growth in monoclonal or most-avid-fraction antibody[9,15,16,22]. (See Table 1 for details.) *d*, Summary of all known alterations from 1968 to 1977 including laboratory-selected variants[3–16,22]. A three-dimensional (stereo) drawing of the same view of the molecule is found in Fig. 3 of ref. 1.

Table 1 Antigenic drift residues

Strain	Residue	Replacement (from)	Replacement (to)	Probable antigenic site	Structural location and probable effects of replacement
MEM/102/72*	144	Gly	Asp	A	External loop. Substitution is both larger and charged. Change from glycine may restrict ϕ, Ψ angles of main chain, altering the loop shape.
	122	Thr	Asn	A?	External extended chain (β). Substitution is larger, 18 Å from residue 144, and may effect the edge of a contact region centred at the 144 loop.
	155	Thr	Tyr	B	4 Å from residue 193 (VIC) and the C-terminal end of the 187–196 α-helix. Substitution is considerably larger and would project over the 'pocket'.
	188	Asn	Asp	B	N-terminal exterior face of α helix 187–196, 18 Å from residue 155 altered charge.
	275	Asp	Gly	C	Near 52–77 disulphide bond bulge. Asp 275 was H-bonded to Asn 53, an interaction that would be lost by the Gly substitution.
	207	Arg	Lys	D?	Subunit interface near residues 99 and 100 of another subunit. Substitution is conservative but shorter and unbranched.
	242	Val	Ile	D??	At the surface but near residues 205 and 207 of same subunit and 221 and 222 of an adjacent subunit, also near the oligosaccharide chain attached at 165. Substitution is larger.
	226	Leu	Gln	?	External on the front of the 'pocket', 8 Å from residue 137 (A), 8 Å from 186 (B?), 16 Å from 242 (D?). Substitution increases size and changes polarity. Gln could reach Ser 136 or Asn 38 for H-bonds: Found in 29C 'avid-fraction' variant. Reverts in VIC/3/75.
	3	Leu	Phe	None	Near membrane, probably inaccessible to antibodies. Reverts in VIC/3/75.
	327	Gln	Arg	None	100 Å from distal tip. Reverts in VIC/3/75.
	31	Asp	Asn	None	82 Å from distal tip.
	78	Val	Gly	None (E?)	Probably buried by oligosaccharide at 81.
	182	Ile	Val	None	Buried residue near 220.
	110	Ser	Leu	None	Near Arg 261, buried. Reverts in VIC/3/75.
VIC/3/75† (VIC/1/75)‡	(145)	Ser	Asn	A	External loop. Substitution is larger and has more potential for H-bonding. If the Asn interacts with the 135/136 region the loop conformation might be altered (VIC/1/75).
	137	Asn	Ser	A?	External 7 Å from residue 145. Substitution is smaller, possibly near an edge of the A site.
	189	Gln	Lys	B	External face of 187–196 α helix. Substitution is larger and charged.
	193	Ser	Asn	B	External face of 187–196 α helix, 4 Å from 189 and 4 Å from 155. Substitution is larger with more H-bond potential. Possible new interaction with Tyr 155.
	278	Ile	Ser	C	Surface residue in 52–277 disulphide bulge. Substitution is smaller and polar and could H-bond to Asn 54.
	(53)	Asn	Asp	C	External in 52–277 disulphide bulge, next to 275. Introduces charge (VIC/1/75).
	201	Arg	Lys	D?	Packed in the subunit interface against 217 of the other subunit, below 188. Substitution is smaller and less bulky and could effect the fit between subunits or the site 'B' helix or be recognized directly if the subunit interface were exposed by a conformational charge.
	217	Ile	Val	D?	Packed in the subunit interface against 201 of the other subunit. Similar to the 201 change. Under 189 (B).
	174	Phe	Ser	D??	External after surface chain, 169 to 173: Pro, Asn, Asn, Asp, Asn. 14 Å from 242. 169 is next to 242.
	164	Leu	Gln	None	Buried next to 127. Substitution gets larger and polar.
	126	Thr	Asn	None	New oligosaccharide attachment site. Carbohydrate would cover a smooth face. Remote possibility that it could reach 133 or 122 and affect site A.
	327				Same as 1968 AICHI strain.
	3				
	226				
	110				
	63	Asp	Asn	None (E?)	New oligosaccharide attachment site, very near old site at 81.
	83	Thr	Lys	None (E?)	Destroys oligosaccharide attachment site at 81.
TEX/1/77§	146	Gly	Ser	A	External loop. Substitution increases size and potential for H-bonding to 135, 136 region, also limits Φ, Ψ angles of main chain (see MEM/102/72 residue 144).
	54	Asn	Ser	C	External as 53 above, next to 278 which is not yet sequenced in TEX/1/77.
	260	Met	Ile	None	Buried next to 177, near 174. Close to site D?.
ENG/187/70§	228	Ser	Gln	None	In pocket near 226. Substitution is much larger, and does not reappear after this strain.
ENG/42/72§	208	Arg	Gly	D?	Near subunit interface, residue 207. Substitution smaller and lacks charge.
Monoclonally selected variant‖	143	Pro	His, Ser Thr, Leu	A	Loop. All these residue 143 variations were selected with a single antibody. Selection with other antibodies has produced identical sequence changes.
	144	Gly	Asp	A	Loop. (See MEM/102/72 residue 144).
	133	Asn	Lys	A	Selected with the same antibody that 143 Pro–Ser was selected.
	54	Asn	Lys	C	External 52–277 disulphide bulge. Substitution is larger and charged.
	205	Ser	Tyr	D	Buried in the subunit interface near 221 and 220 (See VIC/3/75 residue 201).
Avid-fraction variant¶					
30D	186	Ser	Ile	B or D	Below 189 in B-site helix. Ser probably H-bonded to carbohydrate attached at 165. Ile may move carbohydrate affecting site D or directly affect site B.
30D	220	Arg	Ile	D	Subunit interface residue near 203 and 205 of an adjacent subunit 217, of the same subunit. Arg 220 probably H-bonds to Gln 210 of another subunit—the Ile substitution could not make such an interaction.
34D	201	Arg	Gly	D?	Refer to VIC/3/65 residue 201.
29C	226	Leu	Gln	?	Refer to MEM/102/72 residue 226.

* Refs 4, 5, 6.
† Refs 8, 9.
‡ Refs 12, 13.
§ Ref. 9.
‖ Refs 23, 24.
¶ Refs 19, 20.

A/Memphis/102/72: 275, Asp to Gly[4–6]; A/Victoria/1/75 and A/Victoria/3/75: 53, Asn to Asp and 278, Ile to Ser respectively[9,12,13]. The absence of amino acid sequences for the 275 to 278 region of HA_1 from a number of partial sequences may also mean that this list is incomplete[9]. A substitution at position 54, Asn to Lys, in a variant selected by monoclonal antibody prevents its neutralization[22] and variant 30D also fails to react with the same monoclonal antibody[16]. Again the variable amino acids seem to be anchored in a matrix of conserved residues: 50 Lys, 51 Ile (Leu), 52 Cys, 277 Cys, 281 Cys and 286 Gly. The variable residues at all four positions are in direct contact. In the structure of the 1968 haemaglutinin, Asp 275 is hydrogen-

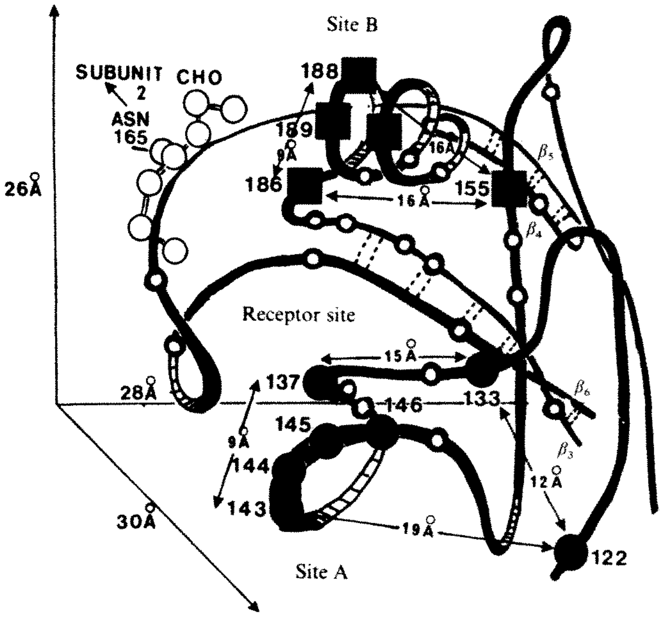


Fig. 3 Schematic drawing showing the size and shape of two of the proposed antigenic sites: ●, A and ■, B (see Fig. 2 and text). Their relationship to the pocket of conserved residues tentatively identified as the host sialic-acid-containing receptor binding site is shown.

bonded to Asn 53, Ile 278 is packed against Asn 54, and Cys 52 bonds to Cys 277. Changes in the types of contact between these residues (for details see Table 1) may be the structural alterations to which antibody binding is sensitive.

Site D: in the first three sites, the amino acid substitutions suggested as causing antigenic variation are external and, in principle, directly recognizable by molecules of the immune system. Site D departs from this apparently simple situation (site C may also). Several amino acid substitutions in the haemagglutinins of both natural and laboratory-selected antigenic variants occur in the interface region between subunits in the haemagglutinin trimer: A/Memphis/102/72: 207, Arg to Lys[4,6]; A/Victoria/3/75: 201, Arg to Lys and 217, Ile to Val[9,12,13]; monoclonal variant: 205, Ser to Tyr[22], and 'avid-fraction' variants 34C: 201, Arg to Gly, and 30D: 220, Arg to Ile[15,16] (see Figs 2, 4). These amino acids may be recognized directly as a result of a relative movement of the globular domains of HA_1 to expose the interface regions. However, it is also possible that the actual antibody binding site is remote from these amino acids but affected by the exact fit at the interface, which might be disturbed by the substitutions listed. A site which spans the monomer–monomer boundary or one near the oligosaccharide chain attached at residue 165, which overlaps that boundary, could be particularly sensitive to alterations in the subunit contact region. Residues 242, Val to Ile and 226, Leu to Gln in A/Memphis/102/72[4,5]; 174, Phe to Ser in A/Victoria/3/75[9,12,13,25] and 226, Leu to Gln (also near sites A and B) in laboratory variant 29C (refs 15, 16) are likely components of this area (see Fig. 2). The amino acid sequences of peptides in the 174 and 242 regions are absent from many partial sequences[9].

Other amino acid substitutions in HA_1 and HA_2 found in nature are considered to be antigenically neutral for reasons detailed in Table 1. For example, the substitution at residue 78, Val to Gly, in the A/Memphis/102/72 haemagglutinin, is buried by the oligosaccharide at position 81. Similarly, in A/Victoria/3/75 at residue 83, the Thr to Lys substitution destroys the oligosaccharide attachment site at residue 81, but the Asp to Asn change at 63 creates a new site—this means that the carbohydrate could occupy almost the same position.

The characterization of new variants may provide evidence for additional sites (for example, 122 and 126, site E?) or extend the size of those proposed.

Antibody interactions and neutralization. This analysis has not defined the extent of each site but only amino acid positions within sites that seem sensitive to substitution when examined serologically, Nevertheless, areas of the surface remain, not in the interfaces nor covered by oligosaccharide chains, that seem serologically unimportant. It is not clear why these regions are not immunogenic, although, for example, the concave shape of the pocket of conserved amino acids tentatively identified with receptor binding[1] probably prevents the production of avid antibodies against that region. The proximity of sites A, B and D to this tentative receptor site (Fig. 3) may be important in the mechanism of neutralization.

Table 1 summarizes changes in the size and/or change of the observed substitutions that would result in a decrease in the binding constant of antibody at a site as a result of steric repulsion or unfavourable charge interactions (the burying of an unpaired charge in the antibody–antigen complex). In addition to these mechanisms, the glycine-rich loop in site A protrudes in such a way as to make so few contacts with the rest of the structure that the effect of mutations in it may be amplified by altering the conformation of the entire loop. The more restricted range of torsion angles possible for amino acids replacing glycine and the potential for hydrogen and ionic bonds from the loop to adjacent regions of the molecule possible with polar or charged substitutions (Table 1) are candidates for such an alteration. (X-ray data on isomorphous crystals of singly substituted, monoclonally selected variant haemagglutinins prepared in collaboration with R. Webster are being collected to allow a direct visualization of these altered sites.)

Antigenic variation and disease incidence (1968–1975)

It seems from the above discussion that the haemagglutinins of both major disease-causing viruses of this period had at least one

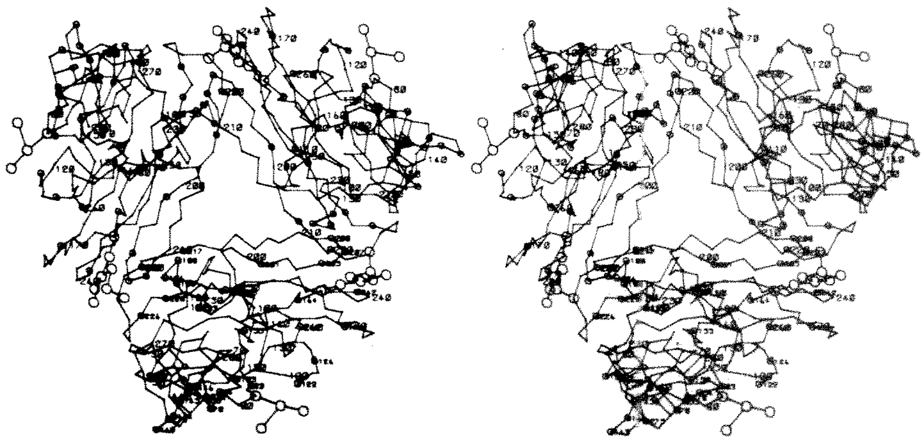


Fig. 4 Stereo drawing of the α-carbon tracing of residues 48–280 (HA_1) of the haemagglutinin trimer. Circled residues are natural and laboratory variations[3–16,22] (see Fig. 2*d*). Tentative (see ref. 1) carbohydrate positions are shown as large open circles for the attachment sites at Asn 81 and Asn 165. The view is down the trimer axis, viewed from the outside towards the membrane.

mutation in each of the four antigenic sites described, and may have required these to avoid virus neutralization. Thus, the viruses acquired the antigenic potential to cause a recurrence of disease in individuals previously infected with viruses of the H3 subtype. Because other characteristics of viruses such as transmissibility and virulence are important in the timing and impact of epidemics, antigenic potential (the recreation of a threshold of susceptible hosts) may, in general, only be a necessary condition for, but not capable of, the generation of epidemics.

The possibility has also been raised[25] that one locus may be more immunogenic than others and therefore, when altered, might be of greater advantage to the virus. Site A, due to its accessibility, seems to be ideally suited both for avid antibody binding and for accepting amino acid substitutions without affecting the main framework of the molecule. To establish the relative importance of the four antigenic sites, a definition will be needed of the spectrum of antibody molecules in human sera with regard to amount, avidity and site specificity.

The character and location of amino acids on the 1968 haemagglutinin structure conserved in three subtypes[3–14] indicates that the basic framework of the molecule is preserved[1]. Because strictly conserved amino acids are found flanking each of the proposed antigenic sites, these structures are also probably present in haemagglutinins from other subtypes. Whether or not the same sites are involved in the interaction of these other haemagglutinins with antibodies is unknown. In the haemagglutinin of other subtypes, these sites have not only varied considerably in amino acid sequence but probable points of oligosaccharide chain attachment are sometimes found within or near regions of the molecule sequentially equivalent to those proposed here as antigenic sites in the Hong Kong haemagglutinins: for example, in fowl plague virus[3] (Hav1) at 133 (site A), 158 (site B) and 240 (possibly site D), and in A/Japan/305/57 (H2) (refs 7, 14) at 169 (site D). It is therefore possible that the immune recognition of such sites in the haemagglutinins of other subtypes is sterically prevented by oligosaccharide chains.

We thank Colin Ward and Theo Dopheide for A/Memphis/102/72 and X:31 A/Aichi/2/68 sequences, Walter Fiers for the sequence of A/Victoria/3/75 and X:31, and Mary Jane Gething and Mike Waterfield for Japan/305/57. We thank Andrew Cherenson for assistance in sequence searches and comparisons, and acknowledge support from NIH grant AI 13654 (D.C.W.). Computing was supported by NSF PC-77-11398.

Received 20 October; accepted 18 November 1980.

1. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289,** 366–373 (1981).
2. Webster, R. G. & Laver, W. G. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 269–314 (Academic, New York, 1975).
3. Porter, A. G. *et al. Nature* **282,** 471–477 (1979).
4. Ward, C. W. & Dopheide, T. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 27–38 (Elsevier, Amsterdam, 1980).
5. Dopheide, T. A. & Ward, C. W. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 21–26 (Elsevier, Amsterdam, 1980).
6. Sleigh, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 69–78 (Elsevier, Amsterdam, 1980).
7. Gething, M. J., Bye, J., Skehel, J. & Waterfield, M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 1–10 (Elsevier, Amsterdam, 1980).
8. Min Jou, W. *et al.* in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 63–68 (Elsevier, Amsterdam, 1980).
10. Verhoeyen, M. *et al. Nature* **286,** 771–776 (1980).
11. Ward, C. W. & Dopheide, T. A. *Biochem. J.* (submitted).
12. Threlfall, G., Barber, C., Carey, N. & Emtage, S. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 51–61 (Elsevier, Amsterdam, 1980).
13. Min Jou, W. *et al. Cell* **19,** 683–696 (1980).
14. Gething, M.-J., Bye, J., Skehel, J. & Waterfield, M. *Nature* **287,** 301–306 (1980).
15. Both, G. W., Sleigh, M. J., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 81–89 (Elsevier, Amsterdam, 1980).
16. Moss, B. A., Underwood, P. A., Bender, V. J. & Whittaker, R. G. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 329–338 (Elsevier, Amsterdam, 1980).
17. Smith, W., Andrewes, G. H. & Laidlaw, P. P. *Lancet* **ii,** 66–68 (1933).
18. World Health Organization (1971) *Bull. Wld Hlth Org.* **45,** 119–124 (1971).
19. Burnett, F. M. *Principles of Animal Virology*, 382 (1955).
20. Pereira, M. S. *Br. med. Bull.* **35,** 9–14 (1979).
21. Stewart-Harris, C. *Br. med. Bull.* **35,** 3–8 (1979).
22. Laver, W. G. *et al.* in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 295–306 (Elsevier, Amsterdam, 1980).
23. Fazekes de St. Groth, S. *Topics Infect. Dis.* **3,** 25–48 (1978).
24. Laver, W. G. *et al. Virology* **98,** 226–237 (1979).
25. Webster, R. G. & Laver, W. G. *Virology* **104,** 139–148 (1980).
26. Atassi, M. Z. & Smith, J. A. *Immunochemistry* **15,** 609–610 (1978).
27. Atassi, M. Z. *Immunochemistry* **15,** 909–936 (1978).
28. Atassi, M. Z. *Immunochemistry* **12,** 423–438 (1975).
29. Public Health Laboratory Service Standing Advisory Committee on Influenza *J. Hyg., Camb.* **78,** 223–233 (1977).
30. Assad, F. A., Cockburn, W. C. & Sundaresan, T. K. *Bull. Wld Hlth Org.* **49,** 219–233 (1973).

A gene chimaera of SV40 and mouse β-globin is transcribed and properly spliced

Gilbert Chu & Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A gene chimaera of the first exon from simian virus 40 (SV40) sequences coding for T antigen placed upstream from the third exon from mouse β-globin, and separated by the resultant new chimaeric intron, was cloned into a bacterial plasmid. When transfected into monkey cells, the gene chimaera was transcribed, polyadenylated and spliced using the donor splice site from SV40 and the acceptor splice site from mouse β-globin. This result suggests that a donor site from one gene can be spliced to an acceptor site from another gene.

THE finding of intervening sequences within mammalian genes has led to speculation that they take part in both phylogeny and ontogeny by facilitating the rearrangement of genes. These ideas are based on the assumption that the cell can properly splice together RNA from any two exons. Thus, evolution and differentiation may proceed in quantum jumps as a new gene is built by rearrangement of exons coding for functional protein domains from different pre-existing genes[1]. Differentiation of the immunoglobulin gene proceeds in such a manner[2]. Similarly, the regulation of a gene may be radically altered if a promoter sequence and short exon are introduced upstream from a series of exons coding for a complete protein. Insertion of this transposable promoter could occur thousands of nucleotides from the protein-coding exons.

If exons can be shuffled arbitrarily and the RNA is properly spliced, then there should be similarity among all donor sites and among all acceptor sites. In fact, the boundaries of intervening sequences are characterized by two consensus sequences, one for donor sites and one for acceptor sites[3]. Furthermore, there are examples of splicing pattern in which two or more donor sites

0028–0836/81/050378—05$01.00

are spliced to a common acceptor site (for example, the early region of SV40)[4,5]. The sequences of such donor sites seem to bear no special relationship other than their relationship to the consensus donor sequence. There are similar examples of splicing patterns in which two or more acceptor sites are spliced to a common donor (for example, the late mRNAs of adenovirus)[6,7]. These observations support, but do not prove, the hypothesis that a donor site from one gene can be spliced to an acceptor site from another gene. The experiment described here tests this hypothesis.

SV40–mouse β-globin chimaera

A gene chimaera of the first exon from SV40 T antigen placed upstream from the third exon from mouse β-globin (MβG), and separated by the resultant new chimaeric intron, was cloned into the plasmid vector pBR322 (see Figs 1, 2). The recombinant genome DNA was isolated from bacteria and transfected into African green monkey kidney cells (CV-1 cell line). RNA was collected from the cells and analysed by a modification of the S_1-endonuclease gel technique[8]—hybridization in 80% formamide to an end-labelled DNA probe, followed by digestion with S_1 endonuclease and electrophoresis on either neutral or denaturing gels.

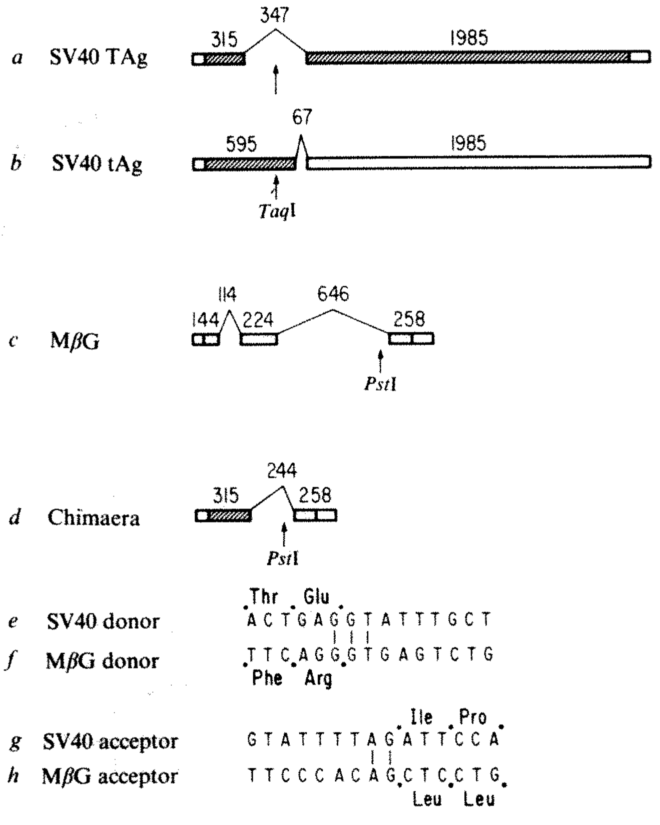


Fig. 1 mRNAs from early SV40, MβG and the gene chimaera. *a*, *b*, The two major SV40 mRNAs produced early in the lytic cycle, corresponding to large T and small t antigens. The coding regions are hatched[4]. *c*, The mouse β-globin mRNA where the coding regions are dotted[15]. *d*, The chimaeric mRNA which would be transcribed from a gene containing SV40 sequences upstream from the SV40 *Taq*I site and containing MβG sequences downstream from the MβG *Pst*I site. The ligation is achieved by introducing *Pst*I linkers (Collaborative Research) to the *Taq*I site in SV40 and annealing to the *Pst*I site in MβG. The chimaeric mRNA contains the first SV40 T-antigen exon and the third MβG exon, with the two exons separated by a chimaeric intron containing both SV40 and MβG sequences. *e*–*h*, The splice junctions from SV40 and MβG. The junction sequences from the SV40 T-antigen intron *e* and *g*[4], and the second MβG intron *f* and *h*[15] are shown. The chimaeric intron consists of SV40 donor and MβG acceptor. Note that the phase of the reading frame relative to the splice junction is the same for the SV40 and MβG introns. Thus the spliced chimaeric mRNA preserves the proper reading frame and permits the translation of a chimaeric protein.

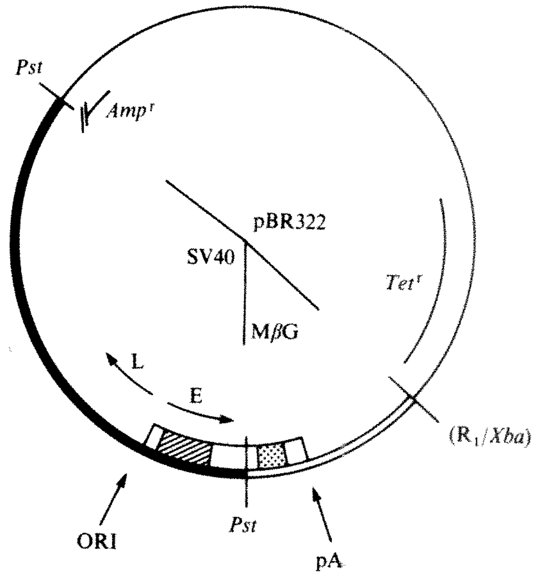


Fig. 2 Structure of the SV40–MβG–pBR322 recombinant genome. SV40 sequences (2,460 base pairs) are indicated by the solid bar, mouse β-globin sequences (~900 base pairs) by the double line and pBR322 plasmid sequences (3,607 base pairs) by the single line. The mouse β-globin (MβG) gene has been sequenced[15]. The SV40 DNA was made from strain 777. Strain 776 has been sequenced[21,22]. Strains 776 and 777 show no difference, by S_1 digestion, in the SV40 early region of interest in this work (G.C., unpublished data). The MβG–pBR322 junction was formed by the blunt-end ligation[23] of the *Xba*I site in the 3′-flanking sequence of MβG with the *Eco*R1 site in pBR322. The SV40–MβG junction was formed by cleavage of SV40 at the *Taq*I site, filling in the staggered end to form a blunt end by Klenow fragment of DNA polymerase I, attachment of *Pst*I linkers to the blunt end, digestion by *Pst*I and staggered-end ligation to the *Pst*I site in the second intron of MβG. The SV40–pBR322 junction was formed by the staggered-end ligation of the *Pst*I site at 0.04 map units in SV40 to the *Pst*I site in the ampicillin-resistance gene of pBR322. The structure was confirmed by analysis with restriction enzymes *Hin*dIII and *Bgl*I, and by double digests with *Hin*dIII and *Pst*I, *Bal*I and *Bgl*I, and *Hin*dIII and *Bal*I. The restriction patterns confirmed expectations from the construction. In particular, the region of the chimaeric intron was shown to contain a *Pst*I site and to be of the length predicted from the SV40 and MβG sequences, to an accuracy of 20 base pairs. The recombinant plasmid was grown in strain C600 of *Escherichia coli* under P2 containment. In SV40, the origin of replication (ORI) and the directions of early (E) and late (L) transcription are indicated. In MβG, the site of polyadenylation (pA) is shown[15]. The DNA sequences coding for the chimaeric mRNA are indicated by the broad bar. The hatched area indicates the coding region of the first exon in SV40 T antigen. The dotted area indicates the coding region of the third exon in MβG. The two coding regions are separated by a chimaeric intron. In pBR322, the ampicillin (*Amp*^{r}) and tetracycline (*Tet*^{r}) resistance genes are indicated by the solid lines. The *Amp*^{r} gene is no longer intact.

Chimaeric mRNA is transcribed, polyadenylated and spliced using the normal MβG acceptor site

The chimaeric genome was transfected into CV-1 cells by a modification of the calcium phosphate co-precipitate technique[9]. It was found that 15% of the cells could be routinely transfected with SV40 DNA in this manner (G.C., unpublished data). To obtain optimal levels of chimaeric mRNA, the chimaeric genome was co-transfected with DNA consisting of SV40 cloned into pBR322 at the *Bam*HI site. The *A* gene product is necessary for the replication of SV40 (ref. 10), and a 10-fold increase in early mRNA is observed between 9 h and 48 h after infection with SV40 virus[11,12]. The SV40–pBR322 recombinant contains an intact early region, and thus replication from the SV40 origin of replication will take place. As the chimaeric genome also contains the SV40 origin of replication, it was hoped that co-transfection would produce a similar increase in chimaeric mRNA, and such an increase was observed. RNA was collected 48 h after transfection. In some cases, poly(A)$^+$ RNA was selected by passage over an oligo(dT) column; all poly(A)$^-$ fractions were treated with DNase I. A DNA probe labelled at the 5′ end (see Fig. 3) was hybridized to the RNA preparations in 80% formamide at 47 °C, treated with S_1 endonuclease and the products electrophoresed.

The results of hybridization of the RNA with the DNA probe labelled at the 5′ end of the *Bal*I site are shown in Fig. 4 for a neutral 2% agarose gel (*a*) and for a denaturing 7 M urea 8% acrylamide gel (*b*). *Bal*I cleaves in the MβG sequences in the chimaera and thus a DNA probe labelled at this site will only hybridize to RNA complementary to globin sequences and not to RNA transcribed from the complementing SV40. The migration of a band on a neutral gel indicates the distance of the 5′ end of the mature RNA from the 5′-end label of the probe. RNA transcribed from the chimaeric genome that was spliced would generate a 415-base pair RNA/DNA hybrid (see Fig. 3). Some 315 and 100 base pairs of this hybrid are contributed by the exons from SV40 and MβG, respectively. Bands, generated by the poly(A)$^+$ fraction from cells transfected with the chimaeric genome after hybridization to DNA labelled at the *Bal*I site, are shown in Fig. 4*a*, lane 4. An RNA/DNA hybrid migrating at a rate expected for a 420 ± 10 base pairs is clearly present. No such band appears in either the poly(A)$^-$ fraction (Fig. 4, lane 5) or a mock transfection sample (Fig. 4, lane 3).

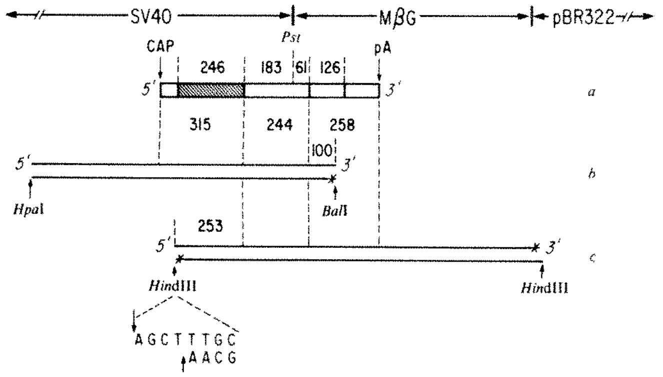


Fig. 3 Gene chimaera and the DNA probes used to analyse the chimaeric mRNA. *a*, The gene chimaera (817 base pairs) is shown from the major initiation site in SV40 (CAP) to the polyadenylation site in MβG (pA). The chimaeric intron is 244 base pairs long, containing 183 base pairs of SV40 sequence and 61 base pairs of MβG sequence. The two coding regions contain 372 base pairs, corresponding to a polypeptide of 124 amino acids. Note that both the coding region and the donor splice site of SV40 t antigen extend beyond the SV40–MβG junction. At the top of the figure the regions corresponding to SV40, MβG and pBR322 sequences are delineated. *b*, the 5′-end-labelled DNA probe was prepared by cutting the recombinant genome with *Bal*I, labelling with [γ-^{32}P]ATP in the presence of t_4 polynucleotide kinase, cutting with *Hpa*I and isolating the fragment by gel elution[24]. *c*, The 3′-end-labelled DNA probe was prepared by cutting the recombinant genome with *Hin*dIII. The DNA was then incubated with dCTP and T_4 DNA polymerase in 10 μl of a buffer of 50 mM Tris *p*H 8, 7 mM $MgCl_2$, 7 mM β-mercaptoethanol at 9 °C to allow the 3′→5′ exonuclease activity of the enzyme to remove nucleotides up to the first position containing a cytosine[25]. After 30 min, [α-^{32}P]dATP was added to allow the 5′→3′ polymerase activity of the enzyme to add labelled nucleotides. After labelling, the fragment was isolated by elution from an agarose gel. The left-hand sequence after *Hin*dIII cleavage is indicated.

The same experiment was done with RNA collected 9 h after transfection with chimaeric DNA alone. About one-tenth the amount of chimaeric RNA was observed in these conditions as compared to co-transfection with SV40. On a neutral gel, the same size band of 420 ± 10 base pairs was observed in both the total cytoplasmic sample and the poly(A)$^+$ fraction, but not in the poly(A)$^-$ fraction or in a mock transfection cytoplasmic sample (data not shown).

The migration of a band on a denaturing gel indicates the distance of the first acceptor splice site from the 5′-end label. Hybridization of DNA labelled at the *Bal*I site to RNA spliced at the acceptor sequences on the third MβG exon would yield a S_1-resistant segment some 100 nucleotides in length. A band migrating at this rate was generated by hybridization of the probe to globin mRNA (Fig. 4*b*, lane 9). A band of the same length, 100 nucleotides, was produced by hybridization to poly(A)$^+$ RNA extracted from a cell transfected with the chimaera. Because a sequencing ladder was run in parallel lanes and because hybridization of the probe to both chimaeric mRNA and MβG RNA produced co-migrating bands, we can assume that the acceptor sites for the chimaeric and MβG splices are identical, 100 nucleotides upstream from the *Bal*I site in the third MβG exon (see Fig. 3).

Chimaeric mRNA is polyadenylated at the normal MβG polyadenylation site and spliced using the normal SV40 donor site

Hybridization results with the probe (labelled at the 3′ end of the *Hin*dIII site in SV40 sequences) are shown in Fig. 5 for a neutral 2% agarose gel (*a*) and for a denaturing 7M urea 8% acrylamide gel (*b*). The migration of a band on a neutral gel indicates the distance of the 3′ end of the mature mRNA from the 3′-end label of the DNA probe. Hybridization of a probe labelled at the 3′ terminus of the *Hin*dIII site to a spliced chimaeric RNA would generate a 511-base pair RNA/DNA hybrid (see Fig. 3). As shown in Fig. 5*a*, lane 4, a band migrating as a 520 ± 20-base pair RNA/DNA hybrid resulted from hybridization to poly(A)$^+$ RNA extracted from cells transfected with the SV40 MβG recombinant.

Because the RNA in this experiment was collected from cells co-transfected with both cloned chimaeric DNA and cloned SV40 DNA, and because the probe is labelled at the 3′ end in the SV40 sequence, there are bands in Fig. 5*a*, lane 4, corresponding to the early SV40 mRNAs for small t and large T antigens. This can be seen by comparing lane 4 with lane 13, which is the result of hybridization to RNA from cells infected with SV40 virus alone. The bands in Fig. 5*a*, lane 4 for small t and large T antigen are much lighter than the band for the chimaeric mRNA because the T_m for the region of SV40 between the origin of replication and *Taq*I sites is ~36 °C (ref. 5) and the T_m for the chimaeric DNA segment is 46 °C. The hybridization was done at 47 °C and therefore strongly favoured the detection of chimaeric mRNA over SV40 early mRNAs.

The migration of a band on a denaturing gel indicates the distance of the first donor splice site from the 3′-end label. The *Hin*dIII site of SV40 is some 256 nucleotides from the donor site in the mRNA for the large T antigen. The denaturing gel in Fig. 5*b* shows that the position of the donor site for the chimaeric splice is at the usual position of the donor for SV40. Part of the same sample analysed on the neutral gel in Fig. 5*a*, lane 4, was loaded on to the denaturing gel in Fig. 5*b*, lane 4. Hybridization of the probe to both chimaeric (Fig. 5*b*, lane 4) and SV40 mRNA (Fig. 5*b*, lane 6) produced bands which co-migrated at a rate consistent with 254 ± 5 base pairs from the donor splice site to the 3′ end of the probe. The markers in lane 1 include a doublet at 276 and 284 base pairs. Comparison of the co-migrating bands with the distance separation of the doublet shows that the donor sites for the chimaeric and SV40 splices are the same, up to an uncertainty of two nucleotides.

Note that Fig. 5*b*, lane 4, is the result of hybridization of the chimaeric probe to RNA from cells co-transfected with both cloned chimaeric and cloned SV40 DNA. When the heteroduplexes are run on a denaturing gel, the bands corresponding to chimaeric mRNA and SV40 large-T mRNA should co-migrate and there will be a second band corresponding to SV40 small-t mRNA. We conclude that the band in Fig. 5*b*, lane 4, is almost entirely due to chimaeric mRNA and only receives a small contribution from SV40 large-T mRNA for two reasons. First, the result in Fig. 5*a*, lane 4, shows that the level of hybridization to chimaeric mRNA is much higher than that for mRNA for either SV40 large T and small-t antigens. Second, the small t and large T mRNAs are present in roughly equal amounts in Fig. 5*a*, lane 4, and the small-t mRNA is repreented by a very faint band in Fig. 5*b*, lane 4.

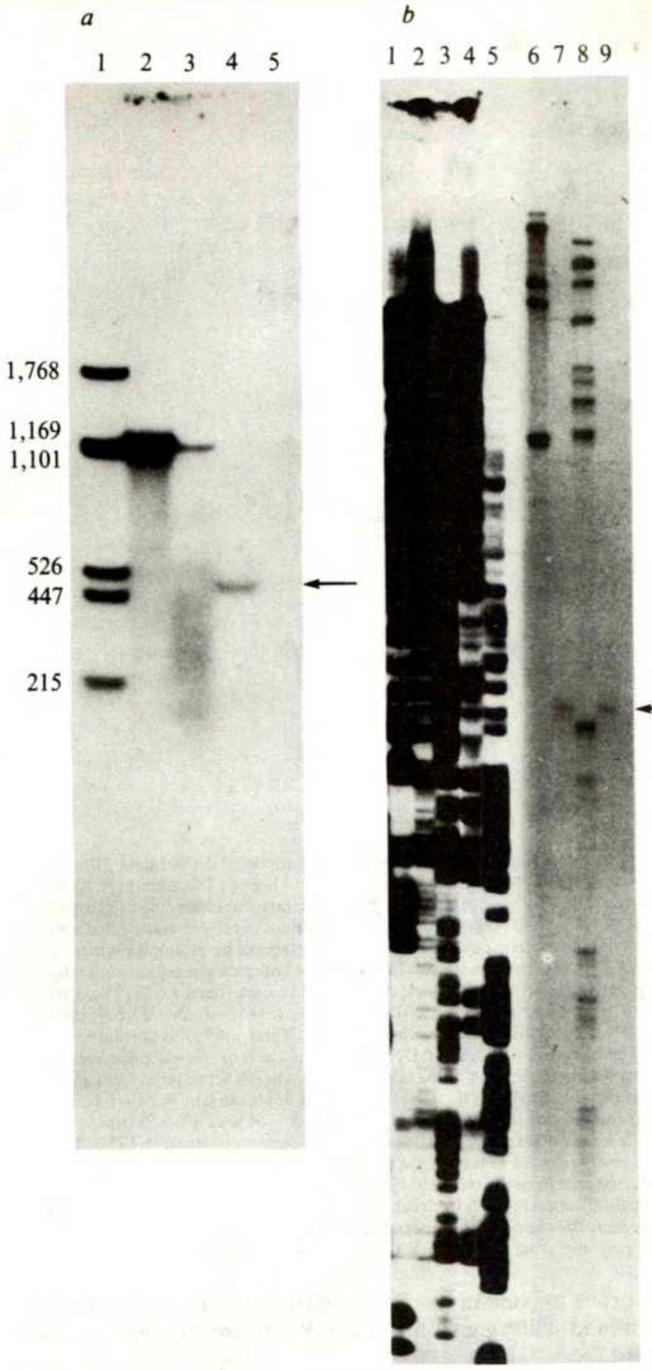


Fig. 4 Analysis with 5′-end-labelled probe of the RNA from cells co-transfected with the gene chimaera and cloned SV40 DNA. The probe is labelled at the *BalI* site of the chimaeric clone with polynucleotide kinase (see Fig. 3*b*) and hybridization was performed in 80% formamide at 47 °C (ref. 8). The transfection procedure is a modification of that used by Graham and van der Eb[9] and Frost and Williams[26]. Here, the cells are exposed to a calcium phosphate/DNA co-precipitate while in suspension (G.C., in preparation). *a*, Neutral 2% agarose gel. Lane 1 shows DNA markers, SV40 DNA cleaved by *Hin*dIII resulting in fragments of length 1,768, 1,169, 1,101, 526, 447 and 215 base pairs. Lane 2 shows intact probe. Lane 3 is the result of hybridization to the cytoplasmic RNA from mock-transfected cells. Lanes 4 and 5 are the poly(A)$^+$ and poly(A)$^-$ cytoplasmic fractions of cells co-transfected with the gene chimaera and cloned SV40 DNA, collected after 48 h. The band in lane 4 corresponds to a heteroduplex of 420 ± 20 base pairs. This value is determined by analysis with heteroduplex markers, as RNA–DNA heteroduplexes migrate slightly slower than DNA–DNA duplexes on a neutral gel (see data in Fig. 5). *b*, 7 M urea 8% acrylamide gel. Lanes 1–5 show a sequencing ladder used as markers. Lane 6 is a marker lane (SV40 cleaved with *Hin*dIII). Lane 7 is the result of hybridization to RNA from co-transfected cells. Lane 8 is a marker lane. Lane 9 is the result of hybridization to MβG mRNA. The bands in lanes 7 and 9 are co-migrating, with an uncertainty of less than one base pair on the sequencing ladder. Because the sequencing ladder in lanes 1–5 contained much more specific activity than the experimental data in lanes 6–9, the gel was photographed to give optimal exposure in lanes 6–9, while preventing overexposure in lanes 1–5.

Analysis of nuclear RNA from co-transfected cells reveals the existence of RNA which is unspliced and which migrates with a size consistent with that expected for the unspliced precursor to the chimaeric message. The RNA observed appears in both poly(A)$^+$ and poly(A)$^-$ fractions and is not observed in the cytoplasm (Fig. 5*a*, lane 11). In addition, two other species of RNA are observed, one which protects the full length of the 5′-end-labelled probe, terminates at the MβG polyadenylation site and is unspliced, and the other which migrates as if it were the spliced product of the first, with the proper chimaeric intron excised (data not shown). These RNA species represent transcripts which must initiate upstream of the *HpaI* site, either in SV40 or pBR322. As before, the spliced mRNA is found only in the poly(A)$^+$ fraction, whereas the unspliced RNA is found in both poly(A)$^+$ and poly(A)$^-$ fractions. Interestingly, neither spliced nor unspliced species accumulate in the cytoplasm.

RNA splicing

We have shown that a gene chimaera of SV40 and MβG is transcribed and spliced with the proper use of the SV40 donor and the MβG acceptor. Because the exons used were obtained from such different sources, monkey virus and mouse, and because the sequences show so little homology between the two donors and between the two acceptors (see Fig. 1), we conclude that splicing can occur over a wide range of donors and acceptors. In the particular case studied here, the spliced mRNA conserves the reading frame (see Fig. 1), and it remains to be seen whether proper splicing will occur if a chimaera is constructed where the spliced mRNA does not conserve the reading frame.

Recent studies of expression of mutated transcription units have suggested the interchangeability among donor sites or among acceptor sites. The mouse κ-chain immunoglobulin gene consists of three exons. In an aberrant κ gene, the donor site at the second exon is deleted and the transcribed message is formed by joining the first and third exons. This message translates into a polypeptide as the two exons have the same reading phase relative to the site of splicing[13]. Similarly, some mRNAs expressed after infection of cells with an Ad2–SV40 hybrid virus, Ad2$^+$ND4, seem to be the product of splicing adenovirus RNA sequences to SV40 sequences[14]. (The term 'chimaeric splice' was used in the title of this article.)

Interchangeability among donor sites and among acceptor sites raises several questions. The two donors from the first and second intervening sequences in MβG show considerable homology to each other[15], and similarly the two acceptors. But the improper splice from the first exon to the third exon has not been observed[16,17]. If in the chimaeric gene the MβG acceptor is spliced to the SV40 donor, which bears little resemblance to the usual MβG donor, why does the improper splice in MβG not occur?

The most attractive explanation is that splicing activity operates in a processive way. That is, it starts at one end of the intervening sequence, for example, at the acceptor site, and completes the splice when it reaches the first donor sequence. This restricts RNA splicing to an intramolecular process and insures against production of a nonfunctional protein by skipping an exon. A paradox immediately arises as there are transcription units with multiple donor sites for a single acceptor site, for example, the two SV40 early messages[4,5]. Also, the adenovirus late mRNAs have multiple acceptor sites for a single donor site[6,7]. A processive splicing mechanism is consistent with such splicing patterns only if splicing activity either (1) reaches the first possible splice termination site and then restarts the splice, or (2) always begins at the donor site in some transcription units and at the acceptor site in others.

The excision of a single intervening sequence may proceed in a processive manner, but it is clear from studies of ovalbumin[18] and vitellogenin nuclear RNA[19], that the excision of a series of intervening sequences does not necessarily proceed in a specific

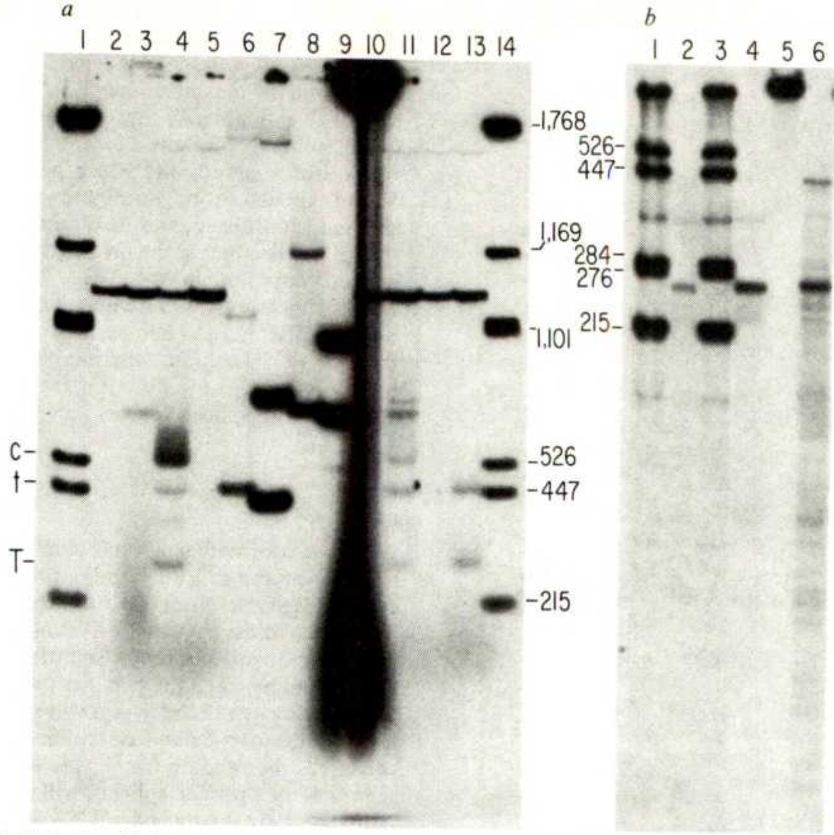


Fig. 5 Analysis with 3′-end-labelled probe of the RNA from cells co-transfected with the gene chimaera and cloned SV40 DNA. The probe is labelled at the *Hin*dIII site of the chimaeric clone with T_4 DNA polymerase (see Fig. 3*c*) and hybridization was done at 47 °C. *a*, Neutral 2% agarose gel. Lane 1 shows DNA markers, SV40 cut by *Hin*dIII. Lane 2 shows intact probe. Lane 3 is cytoplasmic RNA from mock-transfected cells. The band in lane 3 migrating at a rate corresponding to 710 base pairs is an artefact sometimes observed in this series of hybridizations. Lanes 4 and 5 are the poly(A)$^+$ and poly(A)$^-$ cytoplasmic fractions of cells co-transfected with the gene chimaera and cloned SV40 DNA, collected after 48 h. Marked by arrows to the left of the gel are the expected migration patterns of heteroduplexes formed with chimaeric mRNA (c), SV40 small t antigen mRNA (t) and SV40 large T antigen mRNA (T). Note the three bands in lane 4 which migrate at the predicted rates for c, t and T. Lane 10 is nuclear RNA from mock-transfected cells. Lanes 11 and 12 are the poly(A)$^+$ and poly(A)$^-$ nuclear fractions of co-transfected cells. There are bands present corresponding to the artefact in lane 3, and to c, t and T. There is also an extra band migrating at 750 ± 20 base pairs, consistent with the 755 base pairs predicted for an unspliced RNA terminating at the MβG polyadenylation site. Lane 13 is RNA from cells infected with SV40 virus. There are bands present which correspond to mRNAs for t and T. Lane 14 shows the same markers used in lane 1. Lanes 6 and 7 are results with SV40 DNA labelled at the *Bcl*I site and subsequently cleaved with *Pst*I. Lane 7 shows intact DNA prepared in this way and lane 6 shows the result of hybridization of the labelled DNA to cRNA prepared from SV40 DNA with *E. coli* DNA polymerase. Note that the DNA–RNA heteroduplex in lane 6 migrates slightly slower than the DNA–DNA duplex in lane 7. Lanes 8 and 9 show a similar effect to that of lanes 6 and 7, but with SV40 DNA labelled at the *Bcl*I site and subsequently cleaved with *Hin*dIII. *b*, 7 M urea 8% acrylamide gel. Lanes 1 and 3 show DNA markers, SV40 cut by *Hin*dIII, end-labelled by T_4 polynucleotide kinase and recut by *Pst*I, resulting in fragments of length 1,402, 1,101, 911, 526, 447, 284, 276 and 215 base pairs. Lane 2 shows the poly(A)$^+$ nuclear fraction of co-transfected cells from the same hybridization sample loaded on to the neutral gel in Fig. 5*a*, lane 11. Lane 4 shows the poly(A)$^+$ cytoplasmic fraction of co-transfected cells from the same hybridization sample loaded on to the neutral gel in Fig. 5*a*, lane 4. Lane 5 shows the 3′-end-labelled DNA probe. Lane 6 shows the result of hybridization to RNA from cells infected with SV40 virus. (The probe used in lane 6 was the probe used for the hybridization in lanes 2 and 4, but cleaved with *Taq*I. Because the region of SV40 between the *Bgl*I and *Taq*I sites has a T_m of only 36 °C, the hybridization was done at 40 °C.) The bands in lanes 2 and 4 are co-migrating to an accuracy of less than two base pairs.

order. Thus the RNA splicing activity must be capable of interacting in a processive manner in the middle of a long RNA chain containing multiple intervening sequences.

A general model for splicing of pre-messenger RNA would consist of the recognition by a molecular structure of a set of RNA sequences, the consensus sequence, near either the donor or acceptor site, and then translocation of the structure along the intervening sequence until termination occurs at either the acceptor or donor site, respectively. The obvious similarity between this process and translation of mRNA by ribosomes is striking. Perhaps a nuclear analogue of the ribosome, that is, an entity containing both structural RNA and multiple proteins, will prove to contain the RNA splicing activity. The studies of Steitz and colleagues on nuclear RNPs are pertinent in this regard[20].

Andrew Buchman and Paul Berg (personal communication) have constructed a chimaeric gene similar to that described here and have obtained similar results from the study of RNA transcribed from it.

We thank our colleagues for helpful discussions. P.A.S. was supported by American Cancer Society grant MV-37D, NSF grant PCM78-23230 and NIH Program Project grant CA-26717. This work was partially supported by Cancer Center Core grant CA-14051 from the NIH.

Received 26 August; accepted 12 November 1980.

1. Gilbert, W. *Nature* **271,** 501 (1978).
2. Hozumei, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3628–3632 (1976).
3. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4853–4857 (1978).
4. Lebowitz, P. & Weissman, S. M. *Curr. Topics Microbiol. Immun.* **87,** 42–172 (1979).
5. Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1274–1278 (1978).
6. Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74,** 3171–3175 (1977).
7. Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. *Cell* **12,** 1–8 (1977).
8. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732 (1977).
9. Graham, F. L. & van der Eb, A. *J. Virol.* **52,** 456–467 (1973).
10. Tegtmeyer, P. *J. Virol.* **10,** 591–598 (1972).
11. Redd, S. I., Stark, G. R. & Alwine, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3083–3087 (1976).
12. Khoury, G. & May, E. *J. Virol.* **23,** 167–176 (1977).
13. Leder, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 6187–6191 (1978).
14. Khoury, G., Alwine, J., Goldman, N., Gruss, P. & Gilbert, J. *J. Virol.* (in the press).
15. Konkel, D. A., Tilghmann, S. M. & Leder, P. *Cell* **15,** 1125–1132 (1978).
16. Curtis, P. J. & Weissmann, C. *J. molec. Biol.* **106,** 1061–1075 (1976).
17. Ross,, J. *J. molec. biol.* **106,** 403–420 (1976).
18. Nordstrom, J. L., Roop, D. R., Tsai, M.-J. & O'Malley, B. W. *Nature* **278,** 328–331 (1979).
19. Ryffel, G. V., Wyler, T., Muellener, D. B. & Weber, R. *Cell* **19,** 53–61 (1980).
20. Lerner, M. R. & Steitz, J. A. *Proc. natn. Acad. Sci. U.S.A.* **11,** 5495–5499 (1979).
21. Reddy, V. B. *et al. Science* **200,** 494–502 (1978).
22. Fiers, W. *et al. Nature* **273,** 113–120 (1978).
23. Sgaramella, V. *Proc. natn. Acad. Sci. U.S.A.* **69,** 3389–3393 (1972).
24. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
25. Englund, P. T. *J. biol. Chem.* **246,** 3269–3276 (1971).
26. Frost, E. & Williams, J. *Virology* **91,** 39–50 (1978).

LETTERS

Eddy diffusion coefficient for the atmosphere of Venus from radio scintillation measurements

Richard Woo

Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, California 91103

Akira Ishimaru

Department of Electrical Engineering, University of Washington,
Seattle, Washington 98195

Estimates of the vertical mass eddy diffusion coefficient K are necessary to understand the dynamics and photochemistry of a planetary atmosphere. Most previous estimates for the atmosphere of Venus[1–10], pertain to the upper atmosphere. In the lower atmosphere ($\leq$80 km), information on small-scale ($\leq$1 km) temperature fluctuations has been provided by radio scintillations observed during radio occultation by the atmosphere[11–18]. Here we relate the vertical mass eddy diffusion coefficient to these radio scintillation observations so that reliable estimates can be obtained near the tropopause level of 60 km for the first time. Using the recent Pioneer Venus radio occultation measurements[16], and assuming that the turbulence is generated by shear instability, we find that $K = 4\times10^4$ cm^2 s^{-1}, the energy dissipation rate $\varepsilon = 20$ cm^2 s^{-1}, and the dissipation rate of temperature fluctuations $N_\theta = 10^{-3}$ K^2 s^{-1}. These results indicate that small-scale turbulence probably plays a major part in vertical transport near the tropopause of Venus.

Two regions of intense small-scale temperature fluctuations located near 45 and 60 km and coinciding with highly statically stable regions of the atmosphere of Venus have been identified by the radio scintillation measurements[14,15]. We will study the upper region of turbulence. The turbulent eddies are anisotropic and elongated in the horizontal direction, and their wavenumber spectrum is approximately Kolmogorov. The Pioneer Venus orbit 18 entrance radio occultation measurements have been analysed in detail both for the mean temperature profile[19] and for fluctuations in the upper region of turbulence[16] situated above the tropopause. The estimate of structure constant c_T for the temperature fluctuations in the vertical direction is 0.15 K m$^{-1/3}$. This estimate is reliable because: first, the wave propagation theory used to interpret the radio scintillations is the most comprehensive to date[20]; second, contrary to earlier results, the radio scintillations observed by entry probes and orbiting or fly-by spacecraft now show good agreement with each other[21]; and third, temperature fluctuations inferred from the radio measurements agree closely with the Pioneer Venus *in situ* results[16]. The latter were obtained by estimating the deviations from a linear fit to the temperature profile.

For the Kolmogorov spectrum, the structure constant c_T is related to the energy dissipation rate ε and the dissipation rate of temperature fluctuations N_θ by[22,23]

$$c_T^2 = bN_\theta\varepsilon^{-1/3} \tag{1}$$

where b is an empirical constant which is ~2.8 (ref. 24). The eddy diffusion coefficients (diffusivities) for momentum and heat, K_m and K_h respectively, are given by[25–27]

$$\varepsilon = K_m\left(\frac{\partial u}{\partial z}\right)^2(1-R_f) \tag{2}$$

$$N_\theta = K_h\left(\frac{\partial\theta}{\partial z}\right)^2 \tag{3}$$

where R_f is the flux Richardson number, θ is the potential temperature, u is the horizontal wind speed and z is the vertical direction. K_m and K_h are also called eddy viscosity and eddy conductivity respectively. The flux Richardson number is related to the gradient Richardson number Ri by

$$R_f = \frac{K_h}{K_m}Ri \tag{4}$$

and

$$Ri = \frac{g}{\theta}\left[\frac{\partial\theta}{\partial z}\Big/\left(\frac{\partial u}{\partial z}\right)^2\right] \tag{5}$$

where the acceleration due to gravity g is 8.9 m s^{-2} for Venus. The eddy Prandtl number which is K_m/K_h has been found to be close to unity in the Earth's atmosphere[23,28]. If we assume this to be the case of Venus as well and call $K = K_m = K_h$ the eddy diffusion coefficient, we obtain from equations (1)–(5)

$$c_T^2 = b\frac{\theta}{g}\frac{\partial\theta}{\partial z}\varepsilon^{2/3}\frac{Ri}{1-Ri} \tag{6}$$

a result which has been presented by Ottersten[29], and

$$K = \left[\frac{g}{\theta}\left(\frac{1-Ri}{Ri}\right)\right]^{1/2}\frac{c_T^3}{b^{3/2}(\partial\theta/\partial z)^{5/2}} \tag{7}$$

Equation (6) indicates that the scintillations are caused by temperature fluctuations (in the case of Venus, the effect of water vapour fluctuations is negligible) which are directly related to both the mean vertical gradient of potential temperature and ε, a measure of the intensity of the velocity fluctuations. For a neutrally stable atmosphere, even if large velocity fluctuations are present, no scintillations would be detected as $\partial\theta/\partial z = 0$. Thus, although scintillations were not detected in the middle cloud region of the Venus atmosphere near 50 km (ref. 16), it cannot be concluded that mechanical turbulence is absent because the stability of the atmosphere is very close to neutral[30].

The fact that the scintillations are observed in regions of high static stability where high wind shears are also present suggests that the turbulence is produced by wind-shear instability[16]. At 60 km, the Pioneer Venus measurements[30,31] of θ and u yield $Ri = 1.7$ ($\partial\theta/\partial z = 1.6\times10^{-2}$ K m^{-1} and $\partial u/\partial z = 10^{-2}$ s^{-1}). Although Ri is small, it falls short of 0.25 which is the critical value for initiation of dynamic-shear instability indicated by both theoretical and experimental studies[32]. This is, however, not surprising as the Pioneer Venus measurements of temperature and wind velocity lack the spatial resolution to measure the local gradient Richardson number adequately. In the case of the terrestrial free atmosphere, estimates of Ri decrease substantially with decrease in scale over which they are measured[33]. Typically, $Ri = 0.25$ only in thin layers[34] and generally escapes detection unless the spatial resolution is of the order of 100 m. In analysing the scintillation observations of Venus we will therefore assume that $Ri = 0.25$.

Note that when $Ri = 0.25$ and $K_h = K_m$, equations (2)–(5) yield

$$\varepsilon = 3N^2K_h \tag{8}$$

where $N = (g/\theta)\ (\partial\theta/\partial z)^{1/2}$ is the Brunt–Väisälä frequency. Equation (8) is the relationship used by Lilly *et al.*[35] to deduce K_h in the Earth's stratosphere.

Let us now estimate K for Venus at 60 km using equation (7). For $c_T = 0.15$ K m$^{-1/3}$, $b = 2.8$, $Ri = 0.25$, $\partial u/\partial z = 2.6\times10^{-2}$ s^{-1} and $\partial\theta/\partial z = 1.6\times10^{-2}$ K m^{-1}, we obtain $K = 4\times10^4$ cm^2 s^{-1}. In comparing the tropopause regions of the Earth[28,35] and Venus

0028–0836/81/050383—02$01.00

we find that the eddy diffusion coefficients are similar. Note that the assumed wind shear corresponding to $Ri = 0.25$ is $2.6 \times 10^{-2}\ \mathrm{s}^{-1}$, which is a factor of 2.6 higher than the observed value. These strong shears are not unusual in the Earth's free atmosphere and have been observed over altitude intervals of several hundred metres[36].

For $K = 4 \times 10^4\ \mathrm{cm}^2\ \mathrm{s}^{-1}$, equations (2) (3) give $\varepsilon = 20\ \mathrm{cm}^2\ \mathrm{s}^{-3}$ and $N_\theta = 10^{-3}\ \mathrm{K}^2\ \mathrm{s}^{-1}$. The energy dissipation rate ε deduced from the Pioneer Venus radio scintillations can be compared with estimates obtained from the wind measurements. The latter can be obtained using the relationship[22-25]

$$\varepsilon = \frac{(\sigma_v)^3}{L} \tag{9}$$

where σ_v is the r.m.s. velocity and L is the characteristic spatial scale. In the case of the Pioneer Venus measurements[31], we estimate that at 60 km $L \sim 5$ km and $\sigma_v \sim 2\ \mathrm{m\ s}^{-1}$ which is consistent with the Venera 9–12 results[37,38]. This leads to $\varepsilon = 18\ \mathrm{cm}^2\ \mathrm{s}^{-3}$ which agrees well with the radio results. Kerzhanovich *et al.*[38] have analysed the Venera 11 and 12 wind measurements and obtained 50–180 $\mathrm{cm}^2\ \mathrm{s}^{-3}$ for ε above 40 km. All these estimates of ε for Venus fall within the range measured near the Earth's tropopause using *in situ*[28,35] and radar[39,40] observations.

Turbulence in the upper troposphere and lower stratosphere of Venus seems to be similar to that found near the terrestrial tropopause which is due primarily to the Kelvin–Helmholtz instability[41,42]. It is produced in the region of high wind shear[43] which exists above 40 km (ref. 31). Temperature fluctuations are generated in the highly statically stable regions near 45 and 60 km (ref. 30) and these in turn give rise to the radio scintillations that are observed. In the vicinity of 50 km (middle cloud region), the stability of the atmosphere is nearly neutral so that temperature fluctuations are substantially reduced and no scintillations occur. Covey and Schubert[44] have recently discussed the eddy diffusion coefficient for the middle cloud region near the sub-solar point.

Turbulence near the terrestrial tropopause generally takes the form of thin horizontal layers[39-41]. The radio occultation measurements of Venus do not have enough resolution to determine whether the patchiness of terrestrial turbulence also exists in the atmosphere of Venus. However, because the radio scintillation observation is a path-integrated measurement, the estimate of eddy diffusion coefficient obtained in this study is representative of the average atmosphere at 60 km. The equations which were used to estimate K are based on isotropic turbulence and strictly speaking do not apply to the anisotropic turbulence probed by Pioneer Venus. On the other hand, the structure constant c_T used to deduce K is that in the vertical direction and accounts for the fact that the turbulence is stronger in the vertical than in the horizontal direction.

We have found that K is $\sim 4 \times 10^4\ \mathrm{cm}^2\ \mathrm{s}^{-1}$ at 60 km, and believe that this is a good order-of-magnitude estimate. Note that this result is close to the value often assumed in dynamical studies of the atmosphere of Venus[46,47]. Although the scales measured by scintillations are small[16] (≤ 1 km), they are substantially larger than those in the terrestrial troposphere that have been measured and studied extensively with radar[39,40] (≤ 10 m). The fact that the eddy diffusion coefficient estimated from the radio scintillations is consistent with that expected on the basis of vertical cloud-particle distribution[4] and photochemical[10] considerations, implies that small-scale turbulence is probably the dominant mechanism for vertical transport near the tropopause (~ 60 km) in the Venus atmosphere. This is not the case in the Earth's atmosphere[39,48].

Clearly, the radio scintillation technique provides a useful means for determining the eddy diffusion coefficient in regions of the atmosphere where measurements have not yet been made. These radio measurements are applicable to the outer planets, and would yield not only details of the small-scale turbulence in the lower atmosphere, but information on the static stability of the atmosphere as well.

We thank J. W. Armstrong, E. M. Dewan and R. E. Young for useful discussions. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS7-100, sponsored by NASA.

Received 30 September; accepted 26 November 1980.

1. Liu, S. C. & Donahue, T. M. *Icarus* **24**, 148–156 (1975).
2. Sze, N. D. & McElroy, M. B. *Planet. Space Sci.* **23**, 763–786 (1975).
3. Hunten, D. in *Atmospheres of Earth and the Planets* (ed. McCormac, B. M.) 59–72 (Reidel, Dordrecht, Holland, 1975).
4. Prinn, R. *J. atmos. Sci.* **31**, 1691–1697 (1974).
5. Prinn, R. *J. atmos. Sci.* **32**, 1237–1247 (1975).
6. Belton, M. J. S., Smith, G. R., Schubert, G. & Del Genio, A. D. *J. atmos. Sci.* **33**, 1394–1417 (1976).
7. Kumar, S., Hunten, D. M. & Broadfoot, A. L. *Planet. Space Sci.* **26**, 1063–1075 (1978).
8. von Zahn, U., Fricke, K. H., Hoffman, H. J. & Pelka, K. *Geophys. Res. Lett.* **6**, 337–340 (1979).
9. von Zahn, U. *et al. J. geophys. Res.* **85**, 7829–7840 (1980).
10. Winick, J. R. & Stewart, A. I. F. *J. geophys. Res.* **85**, 7849–7860 (1980).
11. Gurvich, A. S. *Doklady. Akad Nauk SSSR* **5**, 1172–1178 (1969).
12. Golitysn, G. S. & Gurvich, A. S. *J. atmos. Sci.* **28**, 138–140 (1971).
13. de Wolf, D. A. *J. geophys. Res.* **76**, 3154–3158 (1971).
14. Woo, R., Ishimaru, A. & Kendall, W. B. *J. atmos. Sci.* **31**, 1698–1706 (1974).
15. Woo, R. *J. atmos. Sci.* **32**, 1084–1090 (1975).
16. Woo, R., Armstrong, J. W. & Ishimaru, A. *J. geophys. Res.* **85**, 8031–8038 (1980).
17. Timofeeva, T. S., Yakovlev, O. I. & Efimov, A. I. *Cosmic Res.* **16**, 285–293 (1978).
18. Kolosov, M. A. *et al. Acta astr.* **7**, 219–234 (1980).
19. Kliore, A. J. & Patel, I. R. *J. geophys. Res.* **85**, 7957–7962 (1980).
20. Woo, R., Ishimaru, A. & Yang, F. C. *Radio Sci.* **15**, 695–703 (1980).
21. Woo, R., Armstrong, J. W.& Kendall, W. B. *Science* **205**, 204–206 (1979).
22. Tatarskii, V. I. *The Effects of the Turbulent Atmosphere on Wave Propagation* (National Technical Information Service, Springfield, 1971).
23. Lumley, J. L. & Panofsky, H. A. *The Structure of Atmospheric Turbulence* (Wiley, New York, 1964).
24. Monin, A. S. & Yaglom, A. J. *Statistical Fluid Mechanics: Mechanics of Turbulence* Vol. 2 (MIT, Cambridge, Massachusetts, 1975).
25. Ishimaru, A. *Wave Propagation and Scattering in Random Media* Vol. 2 (Academic, New York, 1978).
26. Vinnichenko, N. K., Pinus, N. Z., Shmeter, S. M. & Shur, G. N. *Turbulence in the Free Atmosphere* (Consultants Bureau, New York, 1973).
27. Dutton, J. A. & Panofsky, H. A. *Science* **167**, 937–944 (1970).
28. Kennedy, P. J. & Shapiro, M. A. *J. atmos. Sci.* **37**, 986–993 (1980).
29. Ottersten, H. *NOAA Tech. Rep. ERL 243-WPL* (1972).
30. Seiff, A. *et al. J. geophys. Res.* **85**, 7903–7933 (1980).
31. Counselman, C. C. III, Gourevitch, S. A., King, R. W. & Loriot, G. B. *J. geophys. Res.* **85**, 8026–8030 (1980).
32. Woods, J. D. *Radio Sci.* **4**, 1289–1298 (1969).
33. Reiter, E. R. & Lester, P. F. *Arch. Met. Geophys. Bioklim.* **A17**, 1–7 (1968).
34. Bufton, J. W. *J. atmos. Sci.* **30**, 83–87 (1973).
35. Lilly, D. K., Waco, D. E. & Adelfang, S. I. *J. appl. Met.* **13**, 488–493 (1974).
36. Browning, K. A., James, P. K., Parkes, D. M., Rowley, C. & Whyman, A. J. *Nature* **271**, 529–531 (1978).
37. Marov, M. Ya. *A. Rev. Astr. Astrophys.* **16**, 141–169 (1978).
38. Kerzhanovich, V. V., Makarov, Yu. F., Marov, M. Ya., Rozhdestvenkiy, M. K. & Sorokin, V. P. *Academy of Sciences USSR Rep.* (transl. NASA TM-75726, 1979).
39. Crane, R. K. *Radio Sci.* **15**, 177–193 (1980).
40. Gage, K. S., Green, J. L. & van Zandt, T. E. *Radio Sci.* **15**, 407–416 (1980).
41. Dutton, J. A. *Rev. Geophys. Space Phys.* **9**, 613–657 (1971).
42. Thorpe, S. A. *Boundary-Layer Met.* **5**, 95–119 (1973).
43. Crisp, D. & Young, A. T. *Icarus* **35**, 182–188 (1978).
44. Covey, C. & Schubert, G. *Nature* (in the press).
45. Dewan, E. M. *Science* **204**, 832–835 (1979).
46. Goody, R. M. & Robinson, A. R. *Astrophys. J.* **146**, 339–355 (1966).
47. Young, R. E. & Schubert, G. *Planet Space Sci.* **21**, 1563–1580 (1973).
48. Reiter, E. R. *Rev. Geophys. Space Phys.* **13**, 459–474 (1975).

Optical observations of OV−236

D. Wills & Beverley J. Wills

McDonald Observatory, Department of Astronomy,
University of Texas, Austin, Texas 78712

Dent and Balonek[1] recently reported a large outburst in the radio source OV−236 (=1921−293). We have made optical observations of the associated optical object, and report here that it has a high degree of linear polarization, optically variability of ~1 mag in 25 days, and an emission-line spectrum at a redshift $z = 0.3525$. It thus belongs to the small class of extragalactic objects whose properties are intermediate between those of 'classical' QSOs and BL Lac objects.

There has been considerable confusion about the correct optical identification of OV−236. Dent and Balonek[1] refer to the red stellar object suggested in ref. 2, and give an optical

0028-0836/81/050384—03$01.00

position, apparently from ref. 3, but neither is correct. Accurate radio positions that agree within 0.2 arc s have been measured[4,5], but the red object noted in ref. 2 is 37 arc s away, and is therefore definitely excluded as the optical identification. The optical object is correctly identified in ref. 3 and in ref. 5, where its radio and optical position agree within ~1 arc s. We have also measured the 1950.0 optical position from glass plate copies of the Palomar Sky Survey, and find RA = 19 h 21 min 42.26 s, dec = −29°20′27.0″, with r.m.s. uncertainties of ~0.4 arc s; this position is within 0.3 arc s of that given in ref. 5. The optical position given in ref. 3, and repeated in ref. 1, is probably a misprint—there is no object visible there on the Sky Survey plates. The correct identification appears stellar, and about 18 mag, on two pairs of Sky Survey plates (epochs 1954.72 and 1958.53). It may be slightly brighter, and again appears stellar, on the SRC IIIa–J Schmidt plate J2597 (epoch 1976.71), which reaches fainter magnitudes and has better resolution. The object appears to be red on the Sky Survey plates, at least in part because of galactic extinction—its galactic coordinates are $l = 9.3°$, $b = -19.6°$. It has a history of optical variablity[5]. Unfortunately, the radio outburst described in ref. 1 occurred at a time of year when the object was becoming inaccessible to ground-based optical telescopes.

We obtained polarimetry of the object on 12.4 June 1980 UT, using the McDonald Observatory 2.1-m Struve reflector, and the polarimeter described in ref. 6. Visually, the object was estimated to be ~1 mag brighter than on the Sky Survey plates, and equal in brightness to the star 19 arc s following; that is, ~17 mag. The degree of linear polarization was found to be 8.0±0.5% r.m.s., in position angle 149±2°. The galactic reddening that we adopt below implies a contribution of <1% polarization due to the interstellar medium[7].

Spectrophotometry was obtained between 3,900 and 6,900 Å on 14 July 1980 UT, and between 3,400 and 5,700 Å on 8 August 1980, with ~7 Å FWHM resolution, using the McDonald 2.7-m NASA reflector and the instrument described in ref. 8. Figure 1 shows the spectra, calibrated by observations of standard stars to an accuracy of ~10% r.m.s. The difference in continuum levels indicates a fading of nearly 1 mag during the 25 days separating the two sets of observations (J. T. Pollock found a decrease of ~0.7 mag in $m_{\rm pg}$ between 16 July and 7 August).

The first spectrum shows two narrow (~15 Å FWHM) emission lines at 5,042.5 and 6,770.5 Å, which can be unambiguously identified as [O II] and [O III] $\lambda\lambda$3,727, 5,007 at an observed redshift $z = 0.3525 \pm 0.0005$ r.m.s. This spectrum also show an emission line at the expected wavelength of [O III] λ4,959, but no clear sign of Hβ. The second spectrum again shows [O II], at 5,042.9 Å, and a narrow feature at the position of Mg II λ2,798. Although the spectrum is noisy near the position of Mg II, the width of the feature is consistent with those of the forbidden lines. Neither spectrum shows evidence of absorption lines that might indicate the presence of a galaxy.

In the first spectrum the rest-frame equivalent widths of the [O II] and [O III] lines are 3.7 and 9.2 Å, respectively, while in the second one the values for [O II] and Mg II are 9.0 and 4.1 Å, respectively (with r.m.s. uncertainties of typically 10%). The large and significant change in the equivalent width of (O II] is independent evidence for a continuum level change, because the intensity of this line appeared to change by only ~25% between the two observations, a difference that is within the uncertainties in the absolute levels of the two spectra (~10% r.m.s. per observation). The [O II] emission arises in a large low-density region and is not expected to change significantly on a time scale of 25 days (18.5 days at the object). Based on the latter assumption, the fainter spectrum in Fig. 1 has been scaled by a factor of 0.734 so as to equalize the intensities of [O II].

The two spectra in Fig. 1 have the same shapes, with spectral indices $\alpha = 1.3$ and $\alpha = 2.4$ over 5,300–6,750 Å and 3,800–5,300 Å, respectively, after applying a correction for galactic reddening, using $A_{\rm V} = 0.5$ mag (see for example, ref. 10), and defining the spectral index α by $f_\nu \propto \nu^{-\alpha}$. The steepening below 3,600 Å may have been exaggerated by the use of a 2.8 (NS) by 4.5 (EW) arc s aperture and the large zenith distance (~60°) of the object, but we do not believe that these factors contribute to the spectrum curvature between 3,800 and 6,750 Å (this is supported by the good agreement between the spectral shapes in the wavelength region that is common to both spectra). The difference between the two spectra obtained in July and August in the region of overlap is well-represented by a power-law spectrum having $\alpha = 2.1 \pm 0.1$ r.m.s. ($\alpha = 2.9$ before de-reddening); power-law spectra were also derived for the variable components in the variable QSOs 0736+01 and 3C345 (ref. 9).

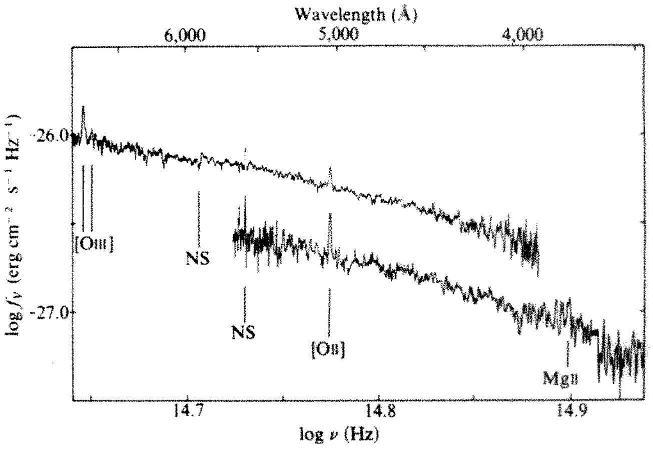


Fig. 1 Spectra of the optical object associated with OV−236, obtained with the McDonald Observatory IDS. The difference in continuum level represents real variability of the object during the 25 days that separate the two observations. Emission lines of Mg II λ2,798, [O II] λ3,727 and [O III] $\lambda\lambda$4,959, 5,007 are marked. NS indicates the positions of incompletely cancelled night sky lines ($\lambda\lambda$5,577, 5,892).

The high degree of linear polarization, the neutral colour and the emission-line spectrum form an unusual, but not unique, combination, placing OV−236 intermediate in properties between the BL Lac objects and QSOs. Most of the objects that are strongly variable at radio and optical frequencies, and that have high (>5%) optical polarization, have optical spectra that show no emission lines when observed with the sensitivity and wavelength resolution appropriate to the present observations. Had the object been 1.5 mag brighter than in July (as it was in mid-1979, ref. 5), with the emission lines retaining their intrinsic strength[9], the lines would have been barely detectable (1–3 Å equivalent width), and we would probably have classified OV−236 as a classical BL Lac object.

Perhaps the object that is most similar to OV−236 is 1400+162 (refs 11, 12), which is neutral in colour, has high optical polarization, and even weaker emission lines than those in OV−236 (from Fig. 2 of ref. 11 we estimate the rest-frame equivalent width of [O II] λ3,727 in 1400+162 to be 1.0 Å, and the intrinsic luminosity of the [O III] λ5,007 line in OV−236 is about 45 times that in 1400+162 (ref. 12)). No radio maps of OV−236 are yet available, but the fact that its 408 MHz flux density[13] is close to the pre-eruption value at high frequencies[1] suggests that the source does not possess extended components with steep radio spectra like those of 1400+162 (ref. 11).

Other possibly similar objects are 0752+258, which has a high degree of linear polarization[14,15], although much stronger emission lines[16,17], and 3C446, with its history of radio and optical variability, its high polarization and its rather weak emission lines[18]. Further observations of this small group of objects may eventually help to elucidate the connection between BL Lac objects and QSOs.

We thank J. T. Pollock for information about the recent optical variability of OV – 236, and the NSF for support.

Received 6 August; accepted 15 November 1980.

1. Dent, W. A. & Balonek, T. J. *Nature* **283,** 747 (1980).
2. Radivich, M. M. & Kraus, J. D. *Astr. J.* **76,** 683 (1971).
3. Peterson, B. A., Bolton, J. G. & Shimmins, A. J. *Astrophys. Lett.* **15,** 109 (1973).
4. Gubbay, J. S. *Astr. J.* **83,** 697 (1978).
5. Donivan, F. F. *et al. Publ. Astrophys. Space Phys.* (in the press).
6. Breger, M. *Astrophys. J.* **233,** 97 (1979).
7. Serkowski, K., Mathewson, D. S. & Ford, V. L. *Astrophys. J.* **196,** 261 (1975).
8. Wills, B. J. & Wills, D. *Astrophys. J.* **238,** 1 (1980).
9. Netzer, H., Wills, B. J., Uomoto, A. K., Rybski, P. M. & Tull, R. G. *Astrophys. J. Lett.* **232,** L155 (1979).
10. Schmidt, M. *Astrophys. J.* **151,** 393 (1968).
11. Baldwin, J. A. *et al. Astrophys. J.* **215,** 408 (1977).
12. Miller, J. S., French, H. B. & Hawley, S. A. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 176 (University of Pittsburgh Press, 1978).
13. Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F. & Sutton, J. M. *Mon. Not. R. astr. Soc.* (in the press).
14. Stockman, H. S. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 149 (University of Pittsburgh Press, 1978).
15. Wills, D., Wills, B. J., Breger, M. & Hsu, J.-C. *Astr. J.* (in the press).
16. Wills, D. & Wills, B. J. *Astrophys. J. Suppl.* **31,** 143 (1976).
17. de Vaucouleurs, G., de Vaucouleurs, A. & Nieto, J.-L. *Astr. J.* **84,** 1811 (1979).
18. Miller, J. S., & French, H. B. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 228 (University of Pittsburgh Press, 1978).

Excited ozone is a possible source of atmospheric N_2O

Sheo S. Prasad

Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, California 91109

In the formation of ozone from O and O_2 in the ground state, the products of the reaction may include electronically excited forms of ozone (O_3^*) which are spectroscopically recognizable in pulsed radiolysis and flash photolysis experiments[1–4]. Indeed, it has been suggested that in the three-body recombination reaction

$$O + O_2 + M \rightarrow O_3^* + M \qquad (1)$$

there is a 62% chance that the ozone product will be in an excited form. Although the identity of these excited precursors of ozone is not known, excited $^3B_2$, $^3A_2$ and $^1A_2$ states of bent ozone and the cyclic ozone have been suggested[5–7], and the existence of electronically excited states of bent ozone below its dissociation limit has been corroborated by large-angle electron scattering experiments[8]. Thus, more than one form of internally excited ozone precursor may be involved. It has been estimated[9] that $\Sigma n(O_3^*) \approx (100-1{,}000)\cdot n(O^1D)$ in the stratosphere. Consequently, even with moderate reactivity, O_3^* may be chemically significant in the stratosphere. We now hypothesize that some of the internally excited ozone, probably that which is electronically excited, may be a potential source of N_2O through the gas-phase reacton (2*a*),

$$O_3^* + N_2 \rightarrow N_2O + O_2 \qquad (2a)$$

$$\rightarrow O_3 + N_2 \qquad (2b)$$

Previously, N_2O was considered[10–12] to form by the reaction

$$O_3 + N_2 \rightarrow N_2O + O_2 \qquad (3)$$

involving unexcited ozone, for which a rate coefficient $k_3 = 5\times10^{-28}\,\mathrm{cm^3\,s^{-1}}$ at 292 K has been reported[13]. This rate constant was estimated by mixing nitrogen and ozone in an absorption tube, measuring ozone disappearance from the absorption in Chappuis band and nitrous oxide formation from the intensity in the 7.8 μm band. However, such a slow reaction between two quite stable molecules is unlikely to proceed without substantial activation energy. In the colder upper atmosphere, therefore, k_3 will probably be much smaller than $5\times10^{-28}\,\mathrm{cm^3\,s^{-1}}$. Hence reaction (3) is not expected to be a significant atmospheric source of N_2O. On the other hand, reaction (2*a*) seems more likely in view of Goody and Walshaw's results[13], because there are many known instances where internal excitation (either vibrational or electronic) in one reactant significantly enhances the reaction.

The wavelength dependence of the quantum yield for N_2O formation, $\phi(N_2O)$, from photolysis of dilute solutions of O_3 in liquid N_2 at 77 K (ref. 14), and the wavelength dependence of $O(^1D)$ quantum yield in gas-phase photolysis of O_3 (refs 15–19) may be interpreted to suggest the possibility of reaction (2*a*). Relevant experimental data are presented in Fig. 1, where $\phi(N_2O)$ obtained by DeMore and Raper[14] is shown with $\phi(O^1D)$ obtained in gas-phase experiments. Over most of the wavelengths used in DeMore and Raper's experiment, $\phi(N_2O)$ is synonymous with $\phi(O^1D)$ in ozone photolysis. Thus, $\phi(N_2O)$ in this experiment begins to decrease rapidly at about 300 nm in agreement with the behaviour of $\phi(O^1D)$ in gas-phase experiments. However, $\phi(N_2O)$ in DeMore and Raper's experiment is perceptible even at 334 nm where $\phi(O^1D)$ is most probably zero according to other experiments. This implies that an additional, low activation energy mechanism of N_2O formation exists, which comes into prominence in DeMore and Raper's experiments at longer wavelengths where $\phi(O^1D)$ vanishes and the reaction

$$O(^1D) + N_2 + M \rightarrow N_2O + M \qquad (4)$$

becomes inoperative. We conjecture that reaction (2*a*) may be this new mechanism. We therefore augmented DeMore and Raper's reaction scheme by including our reactions (1) and (2), and a new expression for N_2O quantum yield is derived:

$$\phi(N_2O/\lambda) \approx \frac{\phi_1(\lambda)}{1 + k_6/k_4} + 0.62\phi_2(\lambda) f_1 f_2 \qquad (5)$$

where 0.62 is the probability of internal excitation in nascent ozone as explained before; f_1 is that fraction of internally excited nascent ozone which could produce N_2O; $f_2 = k_{2a}/k_2$ applicable in their case; ϕ_1 is the quantum yield for $O(^1D)$; ϕ_2 is the same for $O(^3P)$ and k_6 is the rate coefficient for the reaction

$$O(^1D) + N_2 \rightarrow O(^3P) + N_2 \qquad (6)$$

At $\lambda < 300$ nm, $\phi_1 = 1$ and $\phi_2 = 0$, and at 334 nm, $\phi_1 = 0$ and $\phi_2 = 1$, so that $f_1 f_2 \approx 4.2\times10^{-3}$ on the basis of the observed N_2O quantum yields in DeMore and Raper's experiments[14]. On the other hand, failure to detect N_2O, clearly above the background

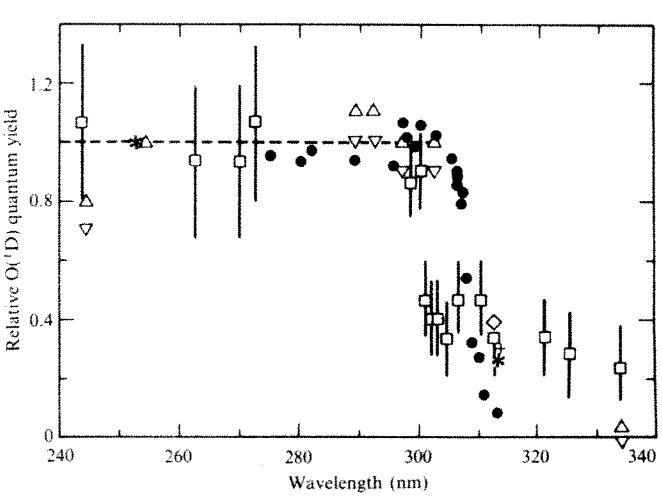


Fig. 1 Relative $O(^1D)$ quantum yields at various wavelengths in the photolysis of ozone. □, Ref. 14; ●, ref. 15; ▽△, ref. 16; +, ref. 17; ◇, ref. 18; *, ref. 19. DeMore and Raper[14] report $\phi(N_2O)$ which is synonymous with $\phi(^1D)$ over most of the wavelengths used by these authors.

0028-0836/81/050386—03$01.00

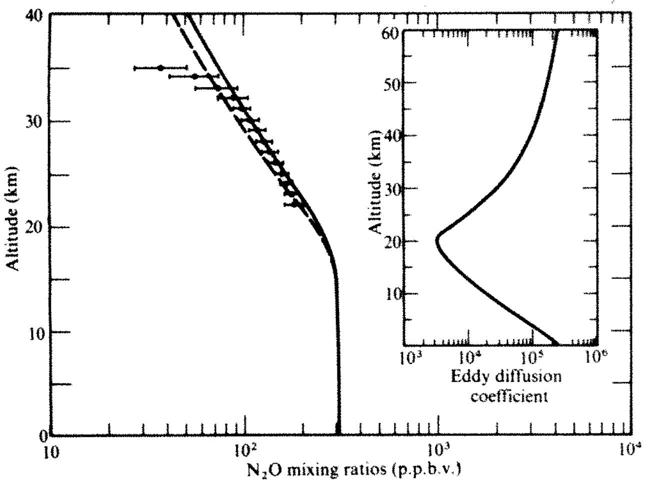


Fig. 2 Theoretical N_2O mixing ratios. Solid line, with the new N_2O source involving excited ozone; dashed line, without this source; |—●—| represent atmospheric measurements of Farmer *et al.*[23]. The inset shows the profile of eddy-diffusion coefficients used in this study.

level, in the experiments of Simonaitis *et al.*[20] places an approximate upper limit of $f_1f_2 \approx 10^{-4}-10^{-5}$ in their experiments conducted at 292 K. Considering widely different experimental conditions and the uncertainties, the difference between the two estimates of f_1f_2 is quite understandable, for example, in terms of possible inverse temperature dependence of f_1f_2. However, because the present study is exploratory we have ignored details, such as the temperature dependence, and have arbitrarily used $f_1f_2 = 10^{-5}$. This choice is consistent with Simonaitis *et al.*'s experiment which was in the gas phase and at a temperature close to the stratospheric temperatures.

The production rate of N_2O through reaction (2*a*) may now be written as:

$$q(N_2O) = \frac{6.2\times10^{-6}k_1n(O)n(O_2)n(M)}{k_2n(M)+1/\tau}k_2n(N_2) \qquad (7)$$

$$\approx \frac{6.2\times10^{-6}(J_{O_3}+2J_{O_2})}{k_2n(M)+1/\tau}k_2n(N_2). \qquad (8)$$

where J_{O_3,O_2} is the number of O_3 (or O_2) photodissociated per second. Approximate equality between reactions (7) and (8) becomes increasingly invalid above about 60 km. But this is immaterial because O_3^*-related source of N_2O is very small at these higher altitudes. As a first approximation, $k_2(=k_{2a}+k_{2b})$ may be assumed to be the same for all internally excited ozone, so that $k_2 \sim (7-0.6)\times10^{-14}\,cm^3\,s^{-1}$ on the basis of several experimental data[1-3]. There are no direct experimental data on the lifetime, τ, of internally excited ozone against non-collisional destruction. We have therefore tried several estimates ranging from a few microseconds to a millisecond. Atmospheric concentrations and fluxes of N_2O corresponding to its production through reaction (2*a*) were then calculated. These calculations were made for the US Standard Atmosphere, 1976, using one-dimensional photochemical-dynamical modelling. We assumed that N_2O molecules were lost by photodissociation and by reaction with $O(^1D)$. $O(^1D)$ concentrations were obtained on the basis of local chemical equilibrium between its production from O_3 photodissociation and loss by quenching collisions with N_2 and O_2. Ozone distribution needed for this and for estimating the solar UV attenuation was taken from US Standard Atmosphere. The empirical method of Shimazaki *et al.*[21] was used to calculate the absorption of solar UV in the O_2 Schuman–Runge bands. We assumed a N_2O mixing ratio of 3.1×10^{-7} at the lower boundary[22] and a zero flux at the upper boundary at 90 km. Calculations were done for a latitude of 30°.

The results in Fig. 2 were obtained by using $\tau = 9\times10^{-6}$ s, $k_2 = 2.0\times10^{-14}\,cm^3\,s^{-1}$ for the excited ozone able to react with N_2 (the eddy-diffusion coefficient, k_{eddy}, profile used is shown as an inset). Our small value of τ is not inconsistent with the experimental data of refs 1–4, because it applies to only a small fraction f_1 of excited ozone which may approach 10^{-5} if f_2 approaches unity. Measurements of the N_2O mixing ratio at 32°N are also shown in Fig. 2 for comparison with our theoretical results. Our theoretical profile using the assumed excited ozone source of N_2O clearly agrees with Farmer *et al.*'s[23] data better than the profile without this source, if we ignore the data points at 34 and 35 km which have significantly larger experimental error bars. This performance of our theory is, of course, dependent on various factors in the calculations, such as the profile of k_{eddy}, and the properties of excited ozone precursors. Nevertheless, with a reasonable choice of k_{eddy}, the assumed N_2O production from excited ozone has reproduced certain observed nitrous oxide mixing ratios as satisfactorily as the theory without this assumed new source. Nitrous oxide mixing ratios exhibit great variability. Naturally, other observed profiles of nitrous oxide mixing ratios may not agree as well with our new theoretical profile. Use of other k_{eddy} profiles, particularly those determined from observed N_2O mixing ratios and the assumption that there is no atmospheric source for this species, will also spoil the good performance of the new theory evident in Fig. 2. These, however, do not contradict the new theory. The present spread in measured N_2O mixing ratios and estimates of k_{eddy} profiles do not allow a conclusive choice between the new theory and the current theory which links atmospheric N_2O solely to microbiological activity on the Earth's surface. Experimental investigations of reaction (2*a*) are needed.

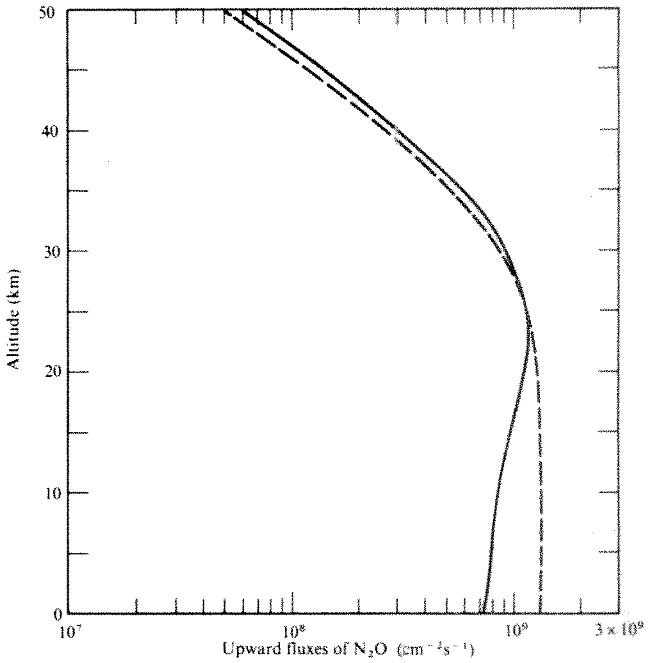


Fig. 3 Upward flux of N_2O corresponding to the profiles of nitrous oxide mixing ratio and eddy-diffusion coefficients shown in Fig. 2.

Fluxes of nitrous oxide in the two limits with and without the new source are presented in Fig. 3. The new source of N_2O implies that the net upward flux of N_2O due to microbiological activities at the Earth's surface may be significantly smaller than the currently accepted value of $(1.3-1.4)\times10^9$ N_2O molecules $cm^{-2}\,s^{-1}$. Thus, this new source tends to strengthen a recent conjecture[24] that the main impact of human beings on the stratospheric N_2O level may not be through the indirect mechanism of increased nitrogen fixation, but through the direct anthropogenic production[25] of N_2O. Besides, the assumed new source of N_2O has two more attributes: (1) it provides a possible new physical basis for latitudinal and temporal variations in N_2O, because upper atmospheric ozone itself has these variations; and (2) this source could impart an additional stability to the ozone layer against external perturbations. Any perturbation which produces an increase in ozone will result in increased N_2O from the new source and thereby in increased NO_x. The increased NO_x will exert a stabilizing effect by arresting further increase in ozone. Similar arguments, but in the opposite sense, would apply to perturbations which produce initial decrease in the ozone layer.

This study represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract no. NAS7-100, sponsored by NASA.

Received 11 August; accepted 28 November 1980.

1. Riley, J. F. & Cahill, R. W. *J. chem. Phys.* **52,** 3297 (1970).
2. Bevan, P. L. T. & Johnson, G. R. A. *JCS Faraday Trans. I.* **69,** 216 (1973).
3. vonRosenberg, C. W. & Trainer, D. W. *J. chem. Phys.* **63,** 5348 (1980).
4. Kleindienst, T., Locker, J. R. & Bair, E. J. *J. Photochem.* **12,** 67 (1980).
5. Wraight, P. C. *Planet. Space Sci.* **25,** 1177 (1977).
6. Hay, P. J., Dunning, T. H. & Goddard, W. A. III *Chem. phys. Lett.* **23,** 457 (1973).
7. Wright, J. S. *Can. J. Chem.* **51,** 139 (1973).
8. Swanson, N. & Cellota, R. J. *Phys. Rev. Lett.* **35,** 783 (1975).
9. Prasad, S. S. & Burton, P. G. *Planet. Space Sci.* **27,** 411 (1979).
10. Bates, D. R. & Witherspoon, A. E. *Mon. Not. R. astr. Soc.* **112,** 101 (1952).
11. Bates, D. R. & Hays, P. B. *Planet. Space Sci.* **15,** 189 (1967).
12. Sinel'nikova, G. Ye. *Atmos. Ocean Phys.* **12,** 106 (1976).
13. Goody, R. M. & Walshaw, C. D. *Q. Jl R. met. Soc.* **79,** 496 (1953).
14. DeMore, W. B. & Raper, O. F. *J. chem. Phys.* **37,** 2048 (1962).
15. Lin, C. L. & DeMore, W. B. *J. Photochem.* **2,** 161 (1973/1974).
16. Jones, I. T. N. & Wayne, R. P. *Proc. R. Soc.* A**319,** 273 (1970).
17. Kuis, S., Simonaitis, R. & Heicklen, J. *J. geophys. Res.* **80,** 1328 (1975).
18. Simonaitis, R., Braslavisky, S., Heicklen, J. & Nicolet, M. *Chem. phys. Lett.* **19,** 601 (1973).
19. DeMore, W. B. & Raper, O. F. *J. chem. Phys.* **44,** 1780 (1966).
20. Simonaitis, R., Lissi, E. & Heicklen, J. *J. geophys. Res.* **77,** 4248 (1972).
21. Shimazaki, T., Ogawa, T. & Farrell, B. C. *NASA Tech. Note TN D-8399* (NASA, Washington DC 1977).
22. *Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport* (US National Academy of Sciences, Washington DC 1979).
23. Farmer, C. B., Raper, O. F., Robbins, B. D., Toth, R. A. & Muller, C. *J. geophys. Res.* **85,** 1621 (1980).
24. Levy, H. II, Mahlman, J. D. & Moxim, W. J. *Geophys. Res. Lett.* **6,** 155 (1979).
25. Weiss, R. F. & Craig, H. *Geophys. Res. Lett.* **3,** 751 (1976).

Flexural strength and porosity of cements

J. D. Birchall, A. J. Howard & K. Kendall

Technical Department, ICI Mond Division, PO Box 8, The Heath, Runcorn, Cheshire WA7 4QD, UK

A curious feature of hydraulic cements, such as those based on calcium silicate, calcium aluminate and calcium sulphate, is that they exhibit similarly low flexural strengths, typically between 3 and 10 MPa, despite their differing chemical composition, varying degrees of hydration and contrasting setting mechanisms[1-3]. Because of these low strength values, unreinforced cements are never used in flexure or tension, and studies of cement strength are usually confined to compression. Those few studies which have considered flexural or tensile failure have concluded that hydraulic cements have an intrinsic maximum tensile strength of about 20 MPa[4,5]. Here we demonstrate that the commonly observed flexural weakness of cement is due to the presence of large voids which are largely undetected by conventional methods of pore analysis such as gas adsorption and mercury porosimetry. The removal of such macro-defects results in flex strengths up to 70 MPa, despite the large volume of gel pores remaining in the material. These strength figures, comparable with those of sintered ceramics, have been achieved without the use of elevated pressures or temperatures, and without fibrous reinforcement.

Much previous work on the fracture of cement has been concerned with compressive failure and with attempts to relate total volume porosity and compressive strength. From such studies, empirical laws such as those of Feret, Abrams and Powers have been derived[6-9]. However, the mode of failure in compression is complex and the relationship of compressive strength to porosity is unlikely to be adequately described by a single equation[10]. On the other hand the tensile and flexural strengths of brittle materials are more amenable to analysis[11-13] and in the Griffith equation,

$$\sigma = (ER/\pi c)^{1/2} \tag{1}$$

tensile cracking stress (σ) is related to an effective critical crack length (c) (which may be a pore); Young's modulus (E), and fracture surface energy (R) being taken as material constants.

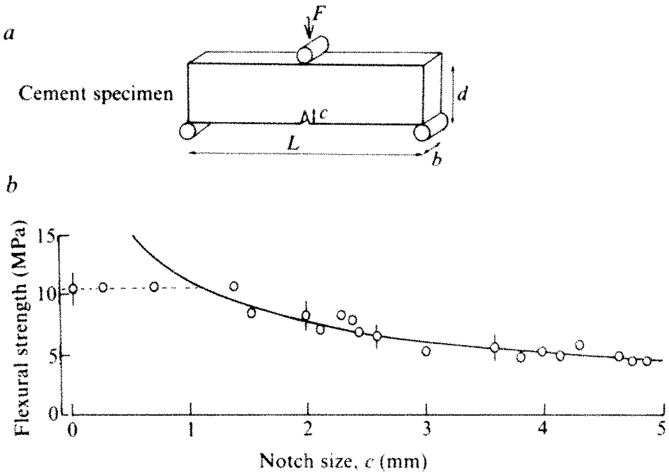


Fig. 1 *a*, The flexural testing arrangement. *b*, Results of flexural testing of notched, Ordinary Portland Cement paste.

The porosity of Portland cement paste is complex and involves a very wide range of pore sizes. Most of the porosity volume is associated with the hydrate gel structure (<100 nm) but larger porosity defects are present. There is a continuing debate as to which microstructural features limit the strength of cement pastes[9]. We have measured the flexural strengths of such pastes and related these through the Griffith equation to the sizes of the largest natural, or artificially-introduced 'pores' within the material. Finally, we have prepared cement paste specimens from which these macroscopic voids have been largely eliminated with a consequent, and dramatic, increase in strength.

The flexural strength testing arrangement is shown in Fig. 1*a*. Bars of cement paste were cast from a slurry containing 100 parts of Ordinary Portland Cement and 22 parts water, cured for 7 days and dried for a further 7 days. In some of the specimens, transverse notches of various depths were cut in the centre of one face using a diamond wheel. Notched and un-notched samples were then bent until cracking occurred. The data are presented in Fig. 1*b* as a Griffith plot of failure stress against

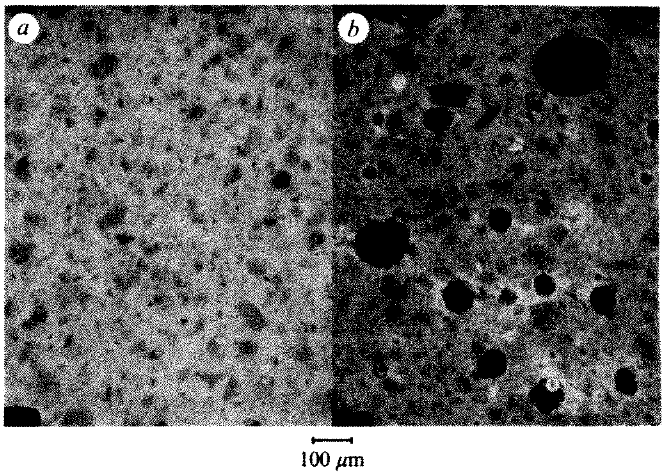


Fig. 2 Optical photomicrographs of polished cement paste sections. *a*, MDF cement. *b*, Ordinary Portland Cement paste. Voids are filled with Indian Ink to improve contrast, and appear black. The grey features are the residues of unhydrated cement grains. ×50.

0028-0836/81/050388—03$01.00

notch depth, the stress σ being calculated from the fracture force F using the relation $\sigma = 1.5\,FL/bd^2$, as the flaw lengths were sufficiently small to avoid significant error[14]. The results gave a reasonable fit to the Griffith curve for notches exceeding 1.1 mm in length. Having calculated E from the flex measurements (20 GPa), the fracture energy $R = 19\ \mathrm{J\,m^{-2}}$ was determined from equation (1). This value, typical of strong, brittle ceramics, indicates that the weakness of cement does not stem from low toughness. The results point to flaw size c as the cause of the weakness. For small notches, the strength was found to be essentially constant at 10.5 ± 2 MPa, the same as un-notched samples. It was inferred that this deviation from the Griffith curve was caused by large natural flaws in the cement.

To investigate this, the porosity of the samples was measured both by mercury porosimetry[15] and by quantitative microscopy[16]. Mercury porosimetry showed the pattern obtained by other workers[17,18], with a median pore diameter of 40–60 nm, a large proportion of gel pores (<7.5 nm diameter) and only 0.1% volume porosity above 15 μm diameter. In contrast, quantitative microscopy of polished sections of the cement demonstrated the presence of significant numbers of large holes, the maximum hole size being measured at 0.83 ± 0.17 mm which agreed satisfactorily with the natural flaw size of 1.1 ± 0.2 mm obtained from the fracture measurements shown in Fig. 1. Furthermore, the volume associated with pores greater than 15 μm diameter was found to be about 5%, far greater than that given by mercury porosimetry, presumably because the mercury method cannot distinguish the pore diameter from the pore entrance diameter, or because the pores are effectively closed[19].

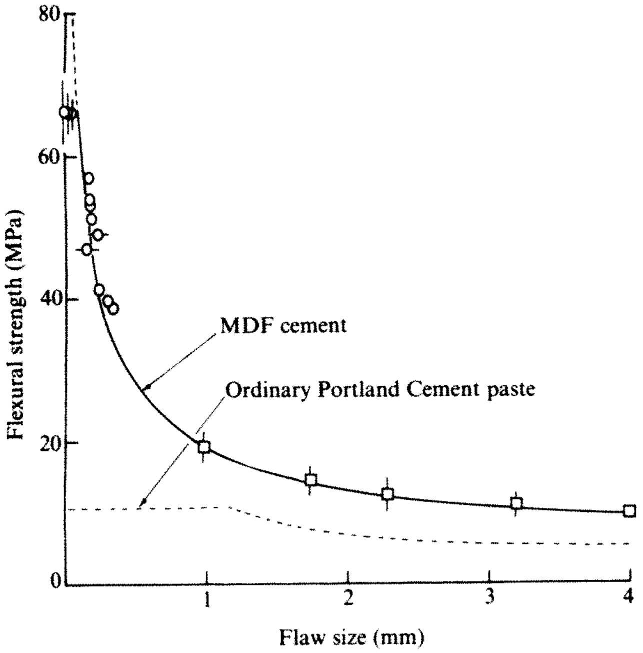


Fig. 3 Results of flexural testing of notched MDF cement (solid line) compared with Ordinary Portland Cement paste (dashed line). □, Flaws introduced by diamond saw; ○, flaws introduced by air bubbles or glass micro-balloons.

Samples of cement paste were then prepared[20] in which the size of large included voids was reduced by a combination of rheological control and efficient mixing. Rheology was controlled, for example, by attention to particle size distribution and the use of rheological aids at low water/cement ratios (<0.2). Such methods reduced the size of the large voids whilst not necessarily reducing the total porosity volume as calculated from true and apparent density. A polished cross-section of a typical product, shown in Fig. 2a, illustrates the freedom from holes exceeding 15 μm in length, contrasting with the ordinary cement paste shown in Fig. 2b. Such preparations, (macro-defect-free (MDF) cement) were subjected to the bend test as previously described. Macrodefects were artificially re-introduced as indicated in Fig. 3 which presents the results together with the data for the ordinary cement samples.

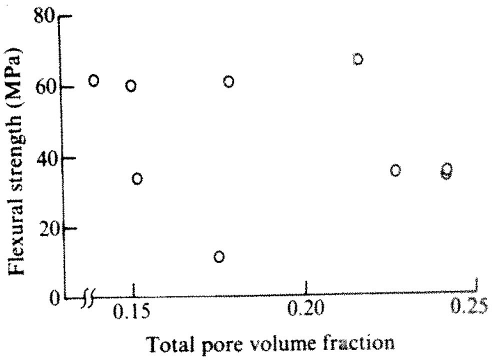


Fig. 4 The lack of correlation of flexural strength of cement paste with total pore volume. (Measured by mercury displacement and helium pycnometry.)

The results fitted the Griffith theory reasonably well but were different in two respects from the original cement. First, the strength at any given flaw size was higher. This was explained by the higher elastic modulus of MDF cement (~40 GPa) and its higher fracture surface energy of $30\ \mathrm{J\,m^{-2}}$. The second difference was that MDF cement followed the Griffith curve to much higher stresses and lower flaw sizes than ordinary cement. When the largest visible holes were ≤90 μm in length (as measured by quantitative microscopy) the samples ceased to fit the Griffith curve but gave the same strength (66 ± 5 MPa) as material containing no holes visible in the optical microscope. Presumably, some flaws other than visible holes, perhaps grain interfaces, dominated at this juncture.

The experiments demonstrate that Griffith theory applies adequately to ordinary cement pastes and that the low flexural strengths normally found are associated with macroscopic flaws rather than microscopic or mineralogical features of the gel. It follows that total pore volume measurements will not correlate with strength except by accident, when pore size distribution remains constant over a series of samples. Figure 4 shows the lack of correlation of total porosity and strength in our experiments. Any attempts to relate flexural strength to gel microstructure, or to seek improved strength by altering microstructure will need to recognize the dominant role of large defects arising during preparation and not necessarily detected by conventional porosimetry.

We have shown that, by removing macroscopic flaws during preparation, cement pastes can be made having flexural strengths (at 60–70 MPa) comparable to sintered ceramics, without the need for fibrous reinforcement or high pressure compaction. This principle is not limited to Portland cement but gives similar improvements when aluminous or sulphate cements are used. Indeed, it should apply generally to all hydraulic cements. Nor is the strengthening limited to tensile fracture; compressive strengths are elevated, being typically greater than 200 MPa.

Received 15 October; accepted 15 November 1980.

1. Soroka, I. *Portland Cement Paste and Concrete* (Macmillan, London, 1979).
2. Neville, A. M. *Properties of Concrete*, 400–403 (Pitman, London, 1963).
3. Schiller, K. K. in *Mechanical Properties of Non-Metallic Brittle Materials* (ed Walton, W. H.) 35–49 (Butterworth, London, 1958).
4. Higgins, D. D. & Bailey, J. E. *J. Mater. Sci.* **11**, 1995–2003 (1976).
5. Berger, R. L., Lawrence, F. V. & Young, J. F. *Cem. Concr. Res.* **3**, 497–508 (1973).
6. Feret, R. *Bull. Soc. Encouragement Indust. Nat.* II 1604 (1897).
7. Abrams, D. *Bull. Struct. Mater. Res. Lab. Lewis Inst. Chicago.* No 1 (1918).
8. Powers, T. C. *Proc. 4th int. Congr. Chem. Cement*, Washington DC, 577–613 (US National Bureau of Standards Monogr. 43, 1962).
9. Sereda, P. J., Feldman, R. F. & Ramachandran, V. S. *Proc. 7th int. Congr. Chem. Cement.* Vol. 1, Principal Rep. VI-1 (3–44) (Editions Septima, Paris 1980).
10. Kendall, K. *Proc. R. Soc.* **A361**, 245–263 (1978).
11. Griffith, A. A. *Phil. Trans. R. Soc.* **A 221**, 163 (1920).
12. Knudsen, F. P. *J. Am. ceram. Soc.* **42**, 376–387 (1959).
13. Davidge, R. W. *Mechanical Behaviour of Ceramics* (Cambridge University Press, London, 1979).
14. Brown, W. F. & Srawley, J. E. *ASTM Spec. Tech. Publ.* 410 (ASTM, Philadelphia, 1967).

15. Rootare, H. M. *Perspect. Powder Met.* **5**, 225–252 (1970).
16. De Hoff, R. T. & Rhines, R. N. *Quantitative Microscopy* (McGraw Hill, New York, 1968).
17. Winslow, D. N. & Diamond, S. *J. Mater.* **5**, 564–585 (1970).
18. Diamond, S. & Dolch, W. *J. Colloid Interface Sci.* **38**, 234–244 (1972).
19. Dullien, F. A. L. *Proc. int. Symp. Rilem/IUPAC Pore Structure and Properties of Materials* Vol 1, C173–186 (Academia, Prague, 1974).
20. European Patent Application No. 80301909.0 Imperial Chemical Industries Ltd (1980).

Conductive polymer: reticulate doping with charge-transfer complex

Jeremiasz K. Jeszka, Jacek Ulański & Marian Kryszewski

Polymer Institute, Technical University of Łódź, 90-924 Łódź,
Żwirki 36, Poland

Additives which can form charge-transfer (CT) complexes enhance the photoconductivity of polymers. In contrast, the dark conductivity of polymers containing this type of additive increases only slightly. Several crystalline CT complexes exhibit high dark conductivity. When introduced into polymer matrix these smuld therefore significantly increase its conductivity provided that the CT complexes are ordered, because their high conductivity is related to the strong interactions between donor and acceptor stacks in the crystal. We describe here a new method of introducing microcrystalline CT complexes into polymers. It consists of growing microcrystals during film casting. The system, poly (bisphenol A carbonate)/(PC) with microcrystals of the CT complex of 7,78,8-tetracyanoquinodimethane (TCNQ) with tetrathiotetracene (TTT), reaches the dark conductivity of 3×10^{-2} Ω^{-1} cm^{-1} at room temperature at 1 wt% of the complex. This conductivity is only slightly temperature dependent.

PC is a polymer with very low electrical conductivity but with good mechanical properties. The crystalline complex of TTT and TCNQ of stoichiometry 1:2 has a conductivity of 10^2 Ω^{-1} cm^{-1} at 300 K measured using a no-contact method at 10^{10} Hz (ref. 1). Casting of a 4 wt% solution of PC and appropriate amounts of TTT and TCNQ on *o*-dichlorobenzene on to hot (390 K) glass substrate produces high-conducting polymer films containing microcrystalline precipitations of the CT complex[2].

Micrographs of the film containing 1 wt% of the complex and that of its cross-section are shown in Fig. 1*a* and *b* respectively. Clearly the microcrystallites are not distributed at random but form dendrite-like paths penetrating the polymer matrix. The microcrystallites disappear irreversibly when a sample is annealed above 450 K, they are not observed in samples cast at temperatures higher than 450 K and in both cases no high conductivity is observed.

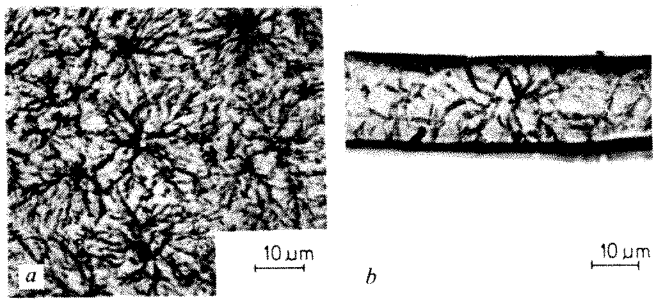


Fig. 1 Micrograph of the PC film containing 1 wt% of the TTT:TCNQ CT complex (*a*) and of its cross-section (*b*). Solutions of PC and TTT and of PC and TCNQ in *o*-dichlorobenzene were prepared separately, and after dissolving the additives at ~420K were mixed together and cast on to glass substrate at 390 K. To preserve appropriate conditions for microcrystallite growth and densification of the polymer matrix, the rate of solvent evaporation has to be controlled. PC (BDH Chemicals, laboratory reagent) was used as received, TTT and TCNQ were purified by recrystallization and the solvent was redistilled before use.

The dependence of the specific conductivity of the films on the content of the CT complex in the sample is shown in Fig. 2. This dependence has a yep-like characteristic and for ~1 wt% of the additive, the conductivity is increased by more than 17 orders of magnitude. When the additive content is higher, a saturation effect is observed. The microscopic investigation of these samples shows that for concentrations below 0.75 wt% the network of microcrystallites does not penetrate all the sample, while for concentrations higher than 1.5 wt% needle-like crystallites of uncomplexed TCNQ and TTT are observed.

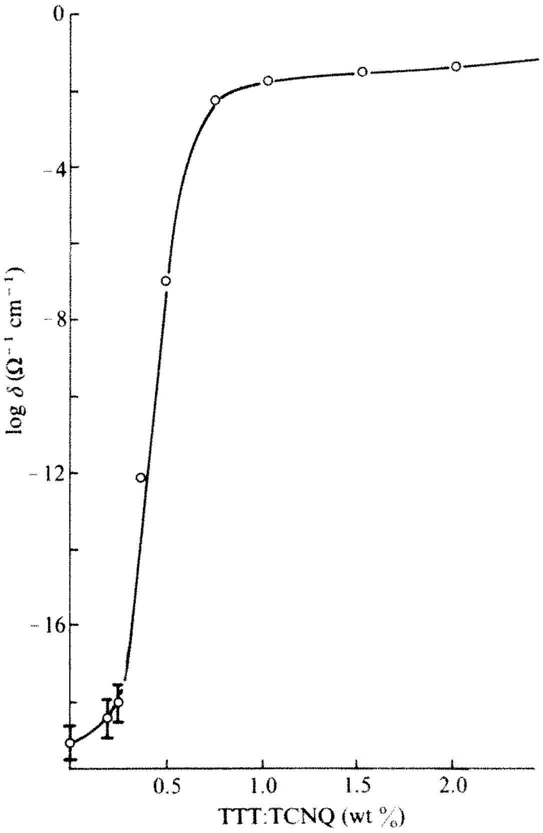


Fig. 2 The dependence of the specific conductivity of the system on the content of the TTT:TCNQ CT complex. Measurements were carried out in a vacuum chamber. For samples having high conductivity (0.5–2.0 wt% of the complex), the four-probe technique was used. Samples having higher resistivity were sandwich-type with evaporated gold electrodes and a guard ring. For samples with intermediate conductivity, both techniques give equivalent results.

The conductivity of homogeneous films obtained by casting at 450 K or by heating conducting samples above 450 K does not exceed 10^{-17} Ω^{-1} cm^{-1} even when the highest concentrations of additives are used. This indicates that the existence of the network of microcrystallites is a prerequisite for the high conductivity of the investigated material.

The temperature dependence of conductivity for samples containing 0.5, 1.0 and 2.0 wt% of TTT:TCNQ is presented in Fig. 3. Contrary to the metal-like temperature dependence of the electrical conductivity found for good conducting CT complexes of TCNQ so-called organic metals, the conductivity of the systems under study shows a weak increase with increasing temperature, diminishing with increasing content of the CT complex.

Current–voltage characteristics of highly conducting samples (concentrations above 0.75 wt%) are ohmic in the investigated temperature range up to the field strength 10^4 V m^{-1}. For higher

0028–0836/81/050390—02$01.00

fields the evolved heat causes overheating of some regions which leads to destruction of the crystals and the samples becoming nonconductive. When the CT complex content is lower, the conductivity seems to be field dependent for the field strength 10^5 V m^{-1} and higher.

To find the real stoichiometry of the complex being formed in these conditions, samples with different donor-to-acceptor molar ratios were investigated. For samples containing 1 wt% of the additives, the molar ratio of TCNQ to TTT changed from 0.45 to 3.0. The conductivity measurements have shown that the highest conductivity is obtained when the molar ratio is 1.0. Also, the microscopic observations of the samples with donor-to-acceptor ratio different from one showed the presence of crystallites characteristic of an excess of uncomplexed component. This suggests that the stoichiometry of the adduct formed within the polymer is 1:1, which differs from the 1:2 stoichiometry of the naked adduct[1].

A relatively high conductivity of the samples containing 1% or more of the CT complex means that the Hall effect may be observed. The experiment was carried out at room temperature in a magnetic field of 10 T for samples containing 1% of additives. The effect was weak, but it could be determined that the excess charge carriers were electrons and their mobility could be estimated to be of the order of 10^{-1} cm^2 V^{-1} s^{-1}.

The step-like dependence of conductivity on the amount of additive present suggests that high conductivity of this material is related to the existence of a microcrystallite network of TTT:TCNQ complex which forms conducting paths, while electrical properties of the polymeric matrix change by only a small extent. On the other hand, the fact that the temperature dependence of conductivity differs from that of highly conducting TCNQ salts indicates that there are some energetic barriers in the paths. These can arise from a separation of some crystallites by the polymer, in which case the increase of conductivity with temperature would be limited by thermal expansion of the polymeric matrix. This effect is, of course, more pronounced for smaller concentrations of the complex (see Fig. 3). It can also explain deviations from Ohm's law in current–voltage characteristics observed for complex concentrations lower than 0.5%.

The value of mobility, which is very high even compared with the mobility in the best polymer photoconductors[3], can be accepted assuming the predominant role of conducting paths and taking into account high values of mobility in organic metals. Assuming one carrier per TTT:TCNQ molecule and full 1:1 complexing, one can calculate for the samples with 1 wt% of additives the density of charge carriers of the order of 10^{19} cm^{-3} and the mobility about 10^{-2} cm^2 V^{-1} s^1, while from the Hall effect we find 10^{18} cm^{-3} and 9×10^{-2} cm^2 V^{-1} s^{-1} respectively which, as the full complexing does not take place, seems to be a reasonable value.

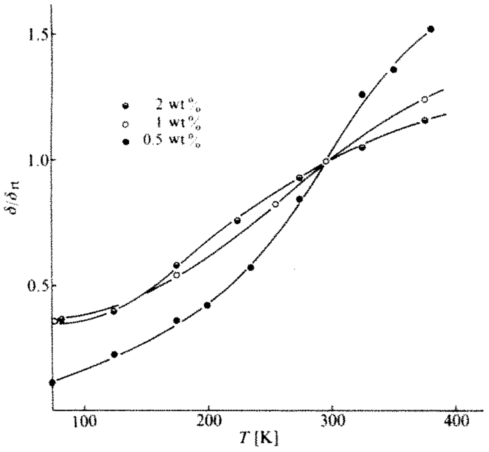


Fig. 3 The temperature dependence of electrical conductivity for the samples with different content of the CT complex. ◒, 2 wt%; ○, 1 wt%; ●, 0.5 wt%

Finally, note that mechanical mixing of polymers with microcrystalline powder of conducting material is much less effective because in this case the amount of conductive additives must usually exceed 20% to achieve comparable values of conductivity[4]. The origin of this difference may be partially explained by a simple geometrical argument: for a needle of length 10 times its other dimensions, the volume is about 50 times smaller than that of a sphere of diameter equal to the needle length. If the needles are perfectly aligned, something of the order of 50 times as much weight of powder would be needed to produce the same effect as oriented needles. In our system, considerable local orientation of microcrystals is obtained by the crystallization of the complex in the polymer matrix in suitable conditions. This orientation is not perfect but a factor of ~10–20 may be expected on the basis of this argument.

The discussed mixtures, due to their high conductivity obtained at low concentration of the CT complex, combine the good application properties of commercial polymers (in contrast with polyacetylenes or $(SN)_x$) with the relatively high conductivity, and form a new class of polymeric materials.

Received 7 August; accepted 14 November 1980.

1. Buravov, L. J. *Soviet Phys. JETP Lett.* **20**, 208–211 (1974).
2. Kryszewski, M., Jeszka, J. K. & Ulański, J. Polish Patent P-215189 (1979).
3. Kryszewski, M. *Semiconducting Polymers* (PWN-Polish Scientific, Warsaw, 1980).
4. Norman, R. H. *Conductive Rubbers and Plastics* (Elsevier, Amsterdam, 1970).

Interlayer formation of humin in smectites

P. Cloos, C. Badot & A. Herbillon

Groupe de Physico-Chimie Minérale et de Catalyse, UCL, and
Section de Physico-Chimie Minérale, MRAC Place Croix du Sud 1,
B-1348 Louvain-la-Neuve, Belgium

The stability of humin, the alkali-resistant part of humic matter in soil[1–4], is usually accounted for by its intimate association with inorganic soil colloids, especially swelling clays and iron compounds. Although the stability of these organo-clay complexes has been explained by interlamellar intercalation, there is evidence[2–4] that large-sized humic acid nuclei are unlikely to penetrate layer silicates in normal soil conditions. The ability of clay minerals to promote polymerization of organic monomers has long been recognized[5], and has been confirmed recently[6–8]; however, the conditions in which these catalytic properties were exhibited were unlike those existing in natural soils. We now report the results of preliminary investigations suggesting that humin-like substances may form between layers of silicate clays from smaller aromatic molecules in conditions that resemble those occurring naturally.

To approach field conditions, we percolated an aqueous aniline (50 p.p.m.) solution through columns containing a 9:1 (w/w) mixture of fine sand and clay. Homoionic Fe(III), Al and Ca exchanged forms of raw montmorillonite from Camp-Berteau (Morocco) were used. An appropriate percolation velocity and clear effluents were obtained by alternating in the columns (height, 140 mm; diameter, 45 mm) ~1-cm thick layers of the sand–montmorillonite mixture and of neat sand, deposited on a 3–4 cm-thick bottom layer of sand supported by glass wool. Apart from interruptions at the weekends, a 50-ml portion of the dilute aniline solution was daily percolated through each column.

During the first treatment, similarly to observations[9] made on 1% clay suspensions, light-green trails and spots developed in the top Fe(III)–clay/sand layers. They turned progressively to green-blue, then to dark greenish-grey whilst increasing in size and extending slowly towards the bottom of the column as more aniline was added. The Al and Ca saturated samples did not

0028-0836/81/050391—03$01.00

Table 1 C and N contents and/or basal spacings at increasing dehydration levels

				Basal spacings $d(001)$ (Å)		
				Air-dry	Degassed (2h) *in vacuo*	
Sample	C (%)	N (%)	C/N	20 °C	20 °C	80 °C
Fe(III)–montmorillonite	—	—	—	15.5	11.8	11.3
Aniline–Fe(III)–montmor	1.12	0.22	6/1	14.2	12.1	11.9
Fe(III)–montmorillonite*	—	—	—	16.6	11.0	10.6
Aniline–Fe(III)–montmor*	1.05	0.37	3.3/1	16.3	12.5	11.3
Fe(III)–montmorillonite†	—	—	—	13.2	10.3	10.3
Aniline–Fe(III)–montmor†	1.13	0.18	7.3/1	13.4	10.65	10.65
DMSO extract‡	38.69	4.74	9.5/1	—	—	—

* Treated with a 0.1 M solution of HCl.

† Treated with a 0.1 M solution of $Na_4P_2O_7$.

‡ After HF attack of the aniline–Fe(III)–montmorillonite complex.

show any change in colour; after a few percolations, the aniline concentration of the effluents equalled that of the treating solution, and washing with distilled water removed most of the organic added. In contrast, no aniline could be detected in the effluents of the column containing Fe(III)–clay, the bottom layer of which had not yet been reached by the green-grey zones at the end of six weeks, when the experiment finished.

After air-drying the column at room temperature, the dark-coloured material was collected and the colloidal organo-clay fraction was separated by sedimentation and freeze-dried. Treating it with a 0.1 M solution of HCl or a 0.1 M alkaline solution of $Na_4P_2O_7$ did not modify either its colour (7.5 GY 4/1 to 10 GY 4/1 on the soil colour chart) or its carbon content (Table 1). Hence, according to the conventional solubility criteria[1–4], the dark-coloured organic substance firmly held by the clay may be termed humin-like material.

Several attempts to extract this humin-like material with various organic solvents failed. Only with dimethyl sulphoxide (DMSO) could we extract part of it from a sample previously attacked with a dilute HF solution. The dark grey to black powder, left after evaporation of the solvent, still had an ash content of 23% for a carbon content of nearly 40% (Table 1).

Assuming interlamellar intercalation, the amount of aniline (Table 1) fixed as a monomolecular layer would cover less than 10% of the internal clay surface[8]. Nevertheless, the basal spacings (Table 1) determined on oriented clay films in a vacuum camera, are systematically higher for the increasingly dehydrated organo-clay complex than for the corresponding neat clay.

A stronger argument for interlamellar intercalation is provided by the IR and ESR spectra (Fig. 1) of the aniline–Fe(III)–montmorillonite complex. The IR spectra, recorded on self-supporting clay films, show evidence of hydration water (ν_{OH} near 3,500, δ_{OH} near 1,640 cm^{-1}), and are characterized by broadening of the aromatic ring vibrations in the 1,600–1,500 cm^{-1} range, by the development of bands near 1,320 and 1,255 cm^{-1} and of a broad charge-transfer absorption extending from roughly 2,000 to beyond 4,000 cm^{-1}. The ESR spectra exhibit a sharp radical signal near the position of a free spinning electron (reference), superimposed on the broad Fe(III) signal. These features are typical for the so-called Type II complexes[10–13], which reportedly form exclusively in the interlayer space of smectites[14]. Their formation generally occurs in dehydrating conditions, through a π interaction between certain aromatic compounds and some transition metal cations occupying exchange sites. This interaction leads to electron transfer resulting in the reduction of the metallic cation and the formation of aromatic radical cations able to participate in polymerization reactions[6–8].

The typical bands of the aniline–Fe(III)–montmorillonite Type II complex persist in the IR spectrum of the DMSO extract (Fig. 1). An additional absorption near 1,665 cm^{-1}, compatible with quinonoid C=O stretching, suggests further oxidation, as do the enhancement of the radical signal (Fig. 1) and the relative decrease in nitrogen content (Table 1) on treating the organo-clay complex with $Na_4P_2O_7$ or DMSO. Similarly, the relative increase in nitrogen content of the acid-treated sample is explainable by partial oxidative hydrolysis, the protonated amino moiety being retained on the clay by exchange wih Fe(III).

Ferric iron was present mostly as hydroxycations ($(Fe(OH)^{2+}$, $Fe(OH)_2^+$) or polymeric species of these, as indicated by the apparent cation exchange capacity of 184 mEquiv. per 100 g, instead of 100 mEquiv. per 100 g when determined on NH_4–montmorillonite at *p*H 7. Ageing for more than 10 months, in aqueous suspension or in ambient conditions, did not prevent Fe(III)–montmorillonite from interacting with aniline as described above. Moreover, aniline seemed to be the first aromatic compound to form a Type II complex on a smectite in aqueous medium, evolving to polymeric species in near-pedogenic conditions.

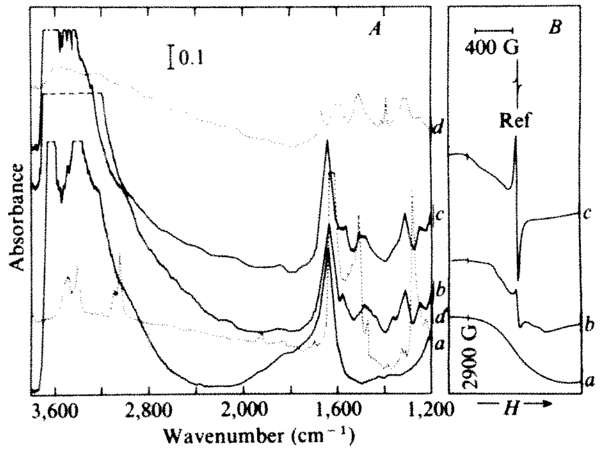


Fig. 1 IR (*A*) and electron spin resonance (*B*) spectra for: *a*, neat Fe(III)–montmorillonite; *b*, aniline–Fe(III)–montmorillonite complex; *c*, same after $Na_4P_2O_7$ treatment. Spectra *a'* and *d* refer to aniline (in CCl_4 solution) and to the DMSO extract of the aniline–Fe(III)–montmorillonite complex (in KBr pellet), respectively.

The above observations demonstrate that interlayer humin formation in smectitic soils or sediments can take place through Type II complexes at the expense of small-sized aromatic molecules. Although aniline is not very representative of biologically generated, natural aromatics—much less than phenolic compounds[1,3], for instance—its behaviour is probably not unique. The important result is the interlamellar polymerization of a simple aromatic monomer, at low concentration in aqueous solution and at low temperature on a smectite containing Fe(III) exchange forms. This type of reaction might well be an alternative for the intercalation of large-sized, polymeric humic acid nuclei, highly improbable in natural conditions.

Received 30 September; accepted 18 November 1980.

1. Schnitzer, M. & Khan, S. V. *Humic Substances in the Environment* (Marcel Decker, New York, 1972).
2. Duchaufour, P. *Pédologie* Vol. 1 (Masson, Paris, 1977).
3. Hayes, M. H. B. & Swift, R. S. in *The Chemistry of Soil Constituents* Ch. 3 (eds Greenland, D. J. & Hayes, M. H. B.) (Wiley, New York, 1978).

4. Theng, B. K. G. *Formation and Properties of Clay-Polymer Complexes* (Elsevier, Amsterdam, 1979).
5. Theng, B. K. G. *The Chemistry of Clay Organic Reactions* (Hilger, London, 1974).
6. Mortland, M. M. & Halloran, L. J. *Am. J. Soil Sci. Soc.* **40,** 367 (1976).
7. Stoessel, F., Guth, J. L. & Wey, R. *Clay Miner.* **12,** 255 (1977).
8. Cloos, P., Moreale, A., Broers, C. & Badot, C. *Clay Miner.* **14,** 307 (1979).
9. Moreale, A. & Van Bladel, R. *Clay Miner.* **14,** 1 (1979).
10. Mortland, M. M. & Pinnavaia, T. J. *Nature phys. Sci.* **229,** 75 (1971).
11. Cloos, P., Vande Poel, D. & Camerlynck, J. P. *Nature phys. Fi.* **243,** 54 (1973).
12. Pinnavaia, T. J., Hall, P. A., Cady, S. S. & Mortland, M. M. *J. phys. Chem.* **78,** 994 (1974).
13. Rupert, P. J. *J. phys. Chem.* **77,** 784 (1973).
14. Doner, H. E. & Mortland, M. M. *Science* **166,** 1406 (1969).

Compositional convection and stratification of Earth's core

David R. Fearn* & David E. Loper

Geophysical Fluid Dynamics Institute, Florida State University,
Tallahassee, Florida 32306

Traditionally, studies[1,2] of the possible stratification of the outer core of the Earth have been based on thermal buoyancy, with the effects of composition being implicitly ignored. However, recent work[3,4] has demonstrated the importance of compositional buoyancy in providing the power necessary to sustain the geodynamo. It seems likely, therefore, that the effects of composition are also important in determining the stratification of the core. We present here a study of the relative importance of thermal and compositional buoyancy in the core. We briefly review the idea that the growth of an iron-rich inner core leads to the release of material which is depleted in iron and therefore buoyant. We then calculate the density profile which results if both this light material and the latent heat released as the inner core grows are redistributed by diffusion. It is found that not only are compositional effects important, they dominate thermal effects in determining the stratification of the core. Finally, using parameter estimates relevant to the core, we show that the diffusive density profile is unstable throughout the outer core except for a layer some 70 km deep adjacent to the mantle–core boundary. We therefore expect that convection fills the entire outer core with the possible exception of a thin layer below the mantle–core boundary.

It has long been believed that the Earth is in thermal equilibrium, with the internal generation of heat by radioactive decay balancing the surface heat flux. Recently, Stacey[5] pointed out that such an equilibrium cannot be maintained because the strength of the radioactivity decays with time and he argued that a significant portion of the surface heat flux is due to the gradual cooling of the entire Earth. This idea is strongly supported by studies[6–8] using parameterized models of thermal convection.

If the Earth is cooling, there will come a time when the core fluid reaches its freezing temperature and solid begins to freeze out. If the freezing temperature gradient is steeper than the actual temperature gradient, then freezing will first occur at the centre of the core, and the frozen solid will accumulate to form a solid inner core[9]. The freezing process results in the release of light material as well as of latent heat. Both are sources of buoyancy and have been proposed[10,11] as the cause of the convective motions necessary for dynamo action.

The core is not a pure material but is composed mainly of iron together with a significant fraction (8–20%) of some lighter material[12]. The precise nature of this material is not important for our purposes and it is sufficient to model the core as a binary alloy of a heavy metal and a light non-metal. We shall assume the mass fraction ξ of light material to be less than that required for an eutectic mixture ($\xi < \xi_e$) as the case $\xi > \xi_e$ proposed by Braginsky[11] has been shown to be geophysically implausible[13]. Metallurgical studies[14] have shown that the solid which freezes out when $\xi < \xi_e$ contains more of the heavy component and is therefore more dense than the surrounding liquid. By conservation of light material, the freezing process must, therefore, leave a light, compositionally buoyant residue in the liquid adjacent to the freezing interface. Thus, as the inner core grows, light material is released and it is this compositional buoyancy which Braginsky[11] proposed as the source of the energy for the geodynamo.

The energetics of Braginsky's proposal have been calculated[3,4] by estimating the rate of release of gravitational potential energy resulting from the net motion of light material upward and heavy material downward, assuming the inner core to have grown at a constant rate over the lifetime of the Earth. This mechanism was found to be capable of supporting a large toroidal magnetic field through dynamo action. Here we study the dynamics of Braginsky's proposal by estimating the relative strengths of the buoyancy forces produced by temperature and compositional differences within the core.

If the inner core is growing at a mass rate $\dot{M}$ and the change in composition on freezing is denoted by ξ_i, Loper and Roberts[15] have shown that the upward flux I of the light material and the thermal diffusive flux Q of heat into the outer core are given by

$$I = \xi_i \dot{M} \quad \text{and} \quad Q = L\dot{M}, \tag{1}$$

where L is the latent heat. The light material must be redistributed throughout the fluid outer core (causing a gradual increase in the mass fraction of light material) and the heat must be transported up to the mantle–core boundary. There are two possible mechanisms for redistribution: diffusion and convection. The mechanism which occurs is determined by the density gradient. A necessary condition for the diffusive state to be unstable to a convective instability is that a parcel of fluid displaced upward isentropically and without change of composition be less dense than its surroundings;

$$(\partial\rho/\partial r)_{S,\xi} < \mathrm{d}\rho_o/\mathrm{d}r \tag{2}$$

where r is the radial coordinate and a subscript o denotes the static diffusive state. Loper and Roberts[15] have shown that for the large length scales in the core, the density difference required to overcome diffusive processes associated with convective motions is very small and to a good approximation equation (2) is also sufficient for convective instability.

Let us adopt equations of state for density and entropy of the form

$$\mathrm{d}\rho = \beta\rho\,\mathrm{d}p - \alpha\rho\,\mathrm{d}T - \bar{\delta}\rho^2\,\mathrm{d}\xi$$

and

$$\mathrm{d}S = -(\alpha/\rho)\,\mathrm{d}p + (C_p/T)\,\mathrm{d}T + \bar{s}\,\mathrm{d}\xi$$

(see ref. 16). We may now rewrite equation (2) as

$$\alpha\rho\,\mathrm{d}T_o/\mathrm{d}r + \bar{\delta}\rho^2\,\mathrm{d}\xi_o/\mathrm{d}r < (\alpha^2 T/C_p)\,\mathrm{d}p/\mathrm{d}r \tag{3}$$

where C_p is the specific heat capacity at constant pressure, α is the thermal expansion coefficient, and $\bar{\delta}$ is the specific volume change with composition.

In the diffusive state the radial fluxes Q_o of heat and I_o of light material are related to the diffusive gradients by

$$I_o = -\rho DA[\mathrm{d}\xi_o/\mathrm{d}r + (\bar{\delta}/\bar{\mu})\,\mathrm{d}p/\mathrm{d}r] \tag{4}$$

and

$$Q_o = -kA\,\mathrm{d}T_o/\mathrm{d}r \tag{5}$$

The constitutive laws (4) and (5) are equivalent to equations (58.11) and (58.12) of Landau and Lifshitz[17] with $Q_o = A\hat{\mathbf{r}}\,.\,[\mathbf{q} - (\bar{\mu} + \bar{s}T)\mathbf{i}]$ and the coefficient k_T set equal to zero. Here $A = 4\pi r^2$ is the area of the sphere of radius r, k is the thermal conductivity, D is the material diffusion coefficient and $\bar{\mu}$ is the energy of mixing. At the inner-core boundary I_o and Q_o are given by equation (1), whereas at the mantle–core boundary we assume there is no flux of light material ($I_o = 0$), and a specified heat flux Q_m. The flux Q_m results from the diffusion of the latent heat $L\dot{M}$ across the outer core, plus the heat capacity released by the cooling of the core and any heat generated by radioactive decay, ohmic heating, and so on. If we assume the internal heat sources to be uniformly distributed throughout the outer core,

---

* Present address: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

0028–0836/81/050393—02$01.00

then

$$Q_o = [(r_m^3 - r^3)L\dot{M} + (r^3 - r_i^3)Q_m]/(r_m^3 - r_i^3) \qquad (6)$$

where r_m is the radius of the outer core and r_i is the radius of the inner core. Similarly, assuming the material flux I_o distributes the light material uniformly throughout the outer core, we have

$$I_o = \xi_i\dot{M}(r_m^3 - r^3)/(r_m^3 - r_i^3). \qquad (7)$$

Combining equations (3)–(7) and assuming a hydrostatic balance, $dp/dr = -\rho g$ we obtain

$$\frac{\alpha^2 T}{C_p} + \frac{(\bar{\delta}\rho)^2}{\bar{\mu}} < \frac{\left[(r_m^3 - r^3)\dot{M}\left(\frac{\alpha L}{k} + \frac{\bar{\delta}\xi_i}{D}\right) + (r^3 - r_i^3)\frac{\alpha}{k}Q_m\right]}{4\pi r^2(r_m^3 - r_i^3)g} \qquad (8)$$

The terms on the left-hand side of equation (8) represent the stabilizing effects of the conduction of heat down the adiabat and the diffusion of material induced by the pressure gradient (barodiffusion), respectively, while the terms on the right-hand side represent the destabilizing effects of the release of latent heat and of light material associated with the freezing process and the heat flux out through the mantle–core boundary. This is a generalization of equation (14) of Loper and Roberts[18].

The instability criterion (8) can be considerably simplified for parameter values thought to be relevant to the core. Loper and Roberts[18] have shown that $\alpha^2 T/C_p \ll (\bar{\delta}\rho)^2/\bar{\mu}$ and $\alpha L/k \ll \bar{\delta}\xi_i/D$ at the inner-core boundary. Using the same sources as Loper and Roberts[18] for the parameter estimates, it can be shown that these inequalities hold good throughout the outer core. We can also obtain a measure of the importance of the term involving Q_m by comparing its maximum value (at $r = r_m$) with the left-hand side of equation (8). Using parameter estimates appropriate at the mantle–core boundary, with $Q_m = 5 \times 10^{12}$ W, we find that $[\alpha Q_m/gk4\pi r_m^2]/[(\bar{\delta}\rho)^2/\bar{\mu}] \simeq 6 \times 10^{-2}$. Thus we may conclude that thermal effects are not of primary importance in determining the stability of the core. To a first approximation, the thermal terms may be neglected and equation (8) reduces to

$$\frac{(\bar{\delta}\rho)^2}{\bar{\mu}} < \frac{(r_m^3 - r^3)\bar{\delta}\xi_i\dot{M}}{4\pi r^2(r_m^3 - r_i^3)gD}. \qquad (9)$$

This is the condition for the diffusive state to be unstable to convective overturning. It depends only on compositional effects and is equivalent to the condition $d\xi_0/dr < 0$.

The diffusive flux of light material has two components (see equation (4)). The first is the familiar flux down the compositional gradient which acts to make the composition homogeneous (minimizing chemical potential energy). If this were the only component, instability would occur throughout the core for any positive growth rate $\dot{M}$. The second component is the barodiffusive flux which acts to separate out the light component from the heavy component (minimizing gravitational potential energy). Together the two fluxes act to redistribute light material in such a way as to minimize the total (gravitational plus chemical) potential energy. If $\dot{M} = 0$ and the net flux of light material in the core is zero, we can see from equation (4) that the compositional gradient $d\xi_o/dr = -(\bar{\delta}/\bar{\mu})\, dp/dr > 0$ is stable. The mass rate of growth $\dot{M}$ must exceed some critical value before the compositional gradient changes sign. As $\dot{M}$ is increased we can see from equation (9) that this happens first at the inner-core boundary ($r = r_i$). As $\dot{M}$ is further increased a layer $r_i \leqq r \leqq r_t < r_m$ develops in which $d\xi_o/dr < 0$. Within this layer, the diffusive solution is unstable and convection occurs.

Loper and Roberts[18] have shown that at $r = r_i$, inequality (9) is satisfied by a wide margin, implying convection in the lower part of the outer core. On the other hand, at $r = r_m$ the right-hand side of equation (9) vanishes so the diffusive density profile must be stable in a layer adjacent to the mantle–core boundary. We determine the depth $h = r_m - r_t$ of this stable layer by finding that value r_t of the radius for which the left- and right-hand sides of equation (9) are equal. Solving for h we find

$$h = \rho^2\bar{\delta}DgV/\xi_i\dot{M}\bar{\mu}, \qquad (10)$$

where $V = 4\pi(r_m^3 - r_i^3)/3$ is the volume of the outer core and we have assumed $h \ll r_m$. Using parameter estimates relevant to the vicinity of the mantle–core boundary (with $D \simeq 6 \times 10^{-9}$ m^2 s^{-1}) we find the depth of the stably stratified layer to be ~70 km. The uncertainties in the estimates of the various parameters (particularly D and $\bar{\mu}$) mean that this figure should be regarded as only an order-of-magnitude estimate of the depth of the stable layer.

A simple calculation equating the kinetic energy of a convecting fluid parcel to the potential energy gained as this parcel penetrates the stable layer suggests that convection will penetrate only a few metres into the stable layer if it is already established. However, if the stable layer were forcibly mixed by some cataclysmic process (such as core formation), the combined action of barodiffusion and the normal diffusive process would act to re-establish the layer, but only very slowly. In fact, it would take a length of time equal to the age of the Earth to establish a layer several tens of kilometres thick by these processes. Further, this process might be severely inhibited if convective motions disrupt the newly forming stable layer before a strongly stabilizing gradient can be established. In view of these caveats, the recent claim[19] that the top of the outer core is stably stratified must be treated with caution.

We thank Paul Roberts for his helpful discussions. This work was supported by the NSF under grant no. EAR78-00818 and is contribution No. 160 of the Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, Florida.

Received 16 September; accepted 27 November 1980.

1. Stacey, F. D. *Physics of the Earth* 2nd edn, Ch. 7 (Wiley, New York, 1977).
2. Jacobs, J. A. *The Earth's Core* Ch. 3 (Academic, London, 1975).
3. Gubbins, D. *J. Geophys.* **43,** 453–464 (1977).
4. Loper, D. E. *Geophys. J. R. astr. Soc.* **54,** 389–404 (1978).
5. Stacey, F. D. *Phys. Earth planet. Inter.* **22,** 89–96 (1980).
6. Sharpe, H. N. & Peltier, W. R. *Geophys. J. R. astr. Soc.* **59,** 171–205 (1979).
7. Schubert, G., Stevenson, D. & Cassen, P. *J. geophys. Res.* **85,** 2531–2538 (1980).
8. Davies, G. F. *J. geophys. Res.* **85,** 2517–2530 (1980).
9. Jacobs, J. A. *Nature* **172,** 297 (1953).
10. Verhoogen, J. *Geophys. J. R. astr. Soc.* **4,** 276–281 (1961).
11. Braginsky, S. I. *Doklady Akad. Nauk SSSR* **149,** 8–10 (1963).
12. Brett, R. *Rev. Geophys. Space Phys.* **14,** 375–383 (1976).
13. Fearn, D. R. & Loper, D. E. *Geophys. Astrophys. Fluid Dyn.* (submitted).
14. Chalmers, B. *Principles of Solidification* (Wiley, New York, 1964).
15. Loper, D. E. & Roberts, P. H. *Geophys. Astrophys. Fluid Dyn.* **16** 83–127 (1980).
16. Loper, D. E. & Roberts, P. H. *Geophys. Astrophys. Fluid Dyn.* **9,** 289–321 (1978).
17. Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* (Pergamon, Oxford, 1959).
18. Loper, D. E. & Roberts, P. H. *Phys. Earth planet. Inter.* (in the press).
19. Whaler, K. A. *Nature* **287,** 528–530 (1980).

Significance of contrasting magmatism in North East Africa and Saudi Arabia

N. B. W. Harris & I. G. Gass

Department of Earth Sciences, The Open University, Walton Hall,
Milton Keynes MK7 6AA, UK

The Red Sea is a north–south rift that split the Arabian–Nubian shield into eastern (Arabian) and western (Egypt and Sudan) segments by seafloor spreading less than 30 Myr ago[1]. Because geological studies in the region have generally have confined by national frontiers, a comparative study of magmatic evolution on either side of the Red Sea has not, so far as we are aware, previously been undertaken, and current models for Arabian–Nubian shield evolution suggest broadly similar histories for the basement on both sides of the Red Sea. However, although the igneous rock geochronology on both sides of the Red Sea indicate that Precambrian magmatic histories were broadly similar, since then they have been quite distinct in that the African crust has been invaded episodically by quartz syenites, nepheline syenites and peralkaline granites between 500 and 80 Myr; no such igneous activity is evident in Arabia (Fig. 1). We propose here that the Red Sea line has had structural significance from at least the end of the Cambrian though whether it defines a line of contrasting upper mantle or lower crust geochemistry, or of contrasting crustal rigidity cannot yet be established.

0028-0836/81/050394—03$01.00

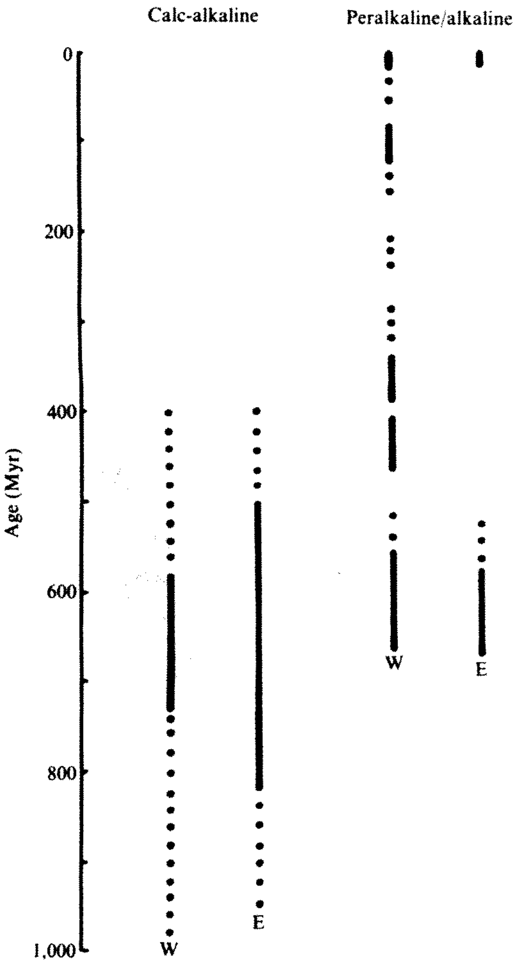


Fig. 1 Age spans of calc-alkaline and peralkaline/alkaline magmatism in the Arabian–Nubian shield. Solid lines from U–Pb zircon or Rb/Sr isochron ($N > 3$) data with $2\sigma < 10\%$. Dotted lines from K/Ar or other Rb/Sr data. E, Eastern shield (Saudi Arabia and Sinai). W, Western shield (Egypt and Sudan). All ages have been recalculated to IUGS recommended constants[2].

Nearly 600 age determinations have been published from the Arabian–Nubian shield, about two-thirds of which are from Saudi Arabia. Although most of these are isolated K/Ar or Rb/Sr ages, and cannot be considered reliable indicators of geological events, available U–Pb zircon and Rb/Sr isochron ages with low errors indicate that on both sides of the Red Sea calc-alkaline igneous activity was more or less continuous between 840 and 500 Myr, being initially dioritic or tonalitic in character until about 700 Myr when granitic magmatism became predominant. There is evidence of sporadic magmatism as early as 980 Myr, and a single three-point isochron of Saudi Arabian metabasalt at 1,165 Myr (ref. 3). Due to daughter isotope leakage, the age of cessation of calc-alkaline magmatism is difficult to define, and some activity may have occurred between 500 and 400 Myr on either side of the Red Sea. However, in contrast to calc-alkaline magmatism, peralkaline and alkaline magmatism have differing age distributions in Africa and Arabia (Fig. 2). In north-west Arabia, excluding Tertiary magmatism associated with Red Sea rifting, peralkaline magmatism seems to be restricted to a geologically brief period[3–6] between 640 and 540 Myr whereas in north-east Africa alkaline intrusions were not only emplaced between 600 and 540 Myr (refs 7, 8), but continued to be so at irregular intervals until 80 Myr ago[8–10]. Many of the available dates for this magmatism are from the K/Ar method, but eight have been confirmed by Rb/Sr isochrons and six more by Rb/Sr mineral ages.

Although the data bank of reliable ages of alkaline intrusions Arabia is not as large as in north-east Africa, there are two independent indications that magmatism did not occur in the eastern shield between 540 and 30 Myr.

(1) In Egypt and north-east Sudan, the mainly Mesozoic Nubian sandstone is intruded by alkaline ring complexes[11,12], whereas peralkaline intrusions in Saudi Arabia are not known to penetrate Phanerozoic sedimentary formations.

(2) In the western shield there is a critical change in magma composition at ~500 Myr. Pre-500 Myr peralkaline granites with smaller volumes of quartz syenites and peralkaline granites were emplaced[7,9,13]. After 500 Myr the dominant rock-type is nepheline syenite with smaller volumes of quartz syenite and peralkaline granite[9,13,14]. This depletion in silica of the younger intrusives indicates a change of either depth or volatile content in the source environment. Analyses from 17 separate alkaline intrusions from Saudi Arabia[15,16] indicate peralkaline granites which are strongly quartz normative which therefore correlate with the pre-50 Myr group from Saudi Arabia. Moreover nepheline-bearing plutons have not been recorded in Arabia so far as we are aware. The absence of silica undersaturated magmas from Saudi Arabia supports the conclusion that the post-500 Myr group from Egypt and Sudan have no equivalent in Saudi Arabia.

Therefore, although there was little if any magmatic activity in Saudi Arabia and Sinai between 500 and 30 Myr, in Egypt and north-east Sudan there were spasmodic and seemingly random high level injections of peralkaline and alkaline magmas throughout this period.

Geochemical considerations of peralkaline granites in general and of Arabian peralkaline granites in particular indicate that such magmas originate by partial melting at depth in the presence of non-hydrous volatiles[16,17]. Of the 16 available Rb/Sr ioschrons of peralkaline and alkaline granites from both sides of the shield, 12 indicate initial $^{87}Sr/^{86}Sr$ ratios between 0.701 and 0.704 implying a source in the upper mantle or lower crust. Four remaining isochrons show ratios between 0.705 and 0.710 which suggest either crustal contamination or a source of high Rb/Sr ratio.

However, as most, if not all, peralkaline magmatism from the shield originates from either the lower crust or upper mantle, its absence in the Arabian segment after 500 Myr may be attributed to: the westward migration of an upper mantle thermal energy source; a more rigid crust to the Arabian segment; or variation in upper mantle geochemistry. These alternatives are now briefly evaluated.

If there has been a westward migration of a high heat-flow region in the upper mantle it could result in an east–west age trend of peralkaline intrusions within the African crust. No such trend is observed and intrusions with age differences of 300–500 Myr are found in juxtaposition[9,13].

Alternatively, it is possible that the Arabian crust became impervious to magmas. Although magma can rise through most of the crust by stokesian diapiric movement, its emplacement into the upper crust is greatly facilitated by the presence of active fractures. It is relevant therefore that fault trends differ in the two segments. In Egypt and the Sudan, an EW–ENE trend is most obvious and controls the emplacment of alkaline intrusions[13], whereas in Arabia the dominant NW–SE Nadj system[18] was active up to 500 Myr ago and is associated with the emplacement of peralkaline granites in Saudi Arabia[15,16]. So, as both segments are cut by faults which were active during the periods of magmatic emplacement, crustal imperviousness to magmatic penetration seems an unlikely explanation for the lack of Phanerozoic peralkaline magmatism in Arabia.

The formation of silica undersaturated alkaline intrusions in the western shield implies the existence of a deep source region low in H_2O and possibly rich in CO_2 (ref. 19). The absence of similar intrusions in the eastern shield may indicate a distinct upper mantle geochemistry and, if this proves to be the case, the distinction existed 450 Myr before the opening of the Red Sea. Although it is possible that the boundary between two upper mantle regions of distinct geochemistry coincidentally become the site of lithospheric rupture, the following model seems to provide a more obvious link between upper mantle heterogeneity and subsequent rifting.

There is strong and increasing evidence[20,21] that the continental crust of the Arabian–Nubian shield was produced by the evolution and accretion of originally intra-oceanic island arc systems during the Upper Proterozoic. This process was complete by about 500 Myr at which time the region attained a continental character, but this was obviously reached in some places before others[16]. The presence of peralkaline granites on both sides of the shield before 500 Myr indicates that the upper mantle or lower crust was able to provide magmas which were emplaced as peralkaline granite melts. Structural and geochemical evidence suggests the Arabian shield was underthrust by one or more Upper Proterozoic, easterly dipping subduction zones[3,21]. In this event, a mantle wedge beneath Arabia and above the subduction zones would be hydrated due to seawater being expelled from the subducted oceanic plate which in time would lower the liquidus temperature of the sub-Arabian peridotite mantle, and the melts produced would be enriched in H_2O and therefore be of calc-alkaline affinity. Peralkaline magmatism, which requires conditions of low P_{H_2O}, would only be initiated after subduction had ceased. However, the sub-Arabian mantle and lower crust would be depleted in components contributing to low temperature melts, and therefore could not provide peralkaline melts after ~500 Myr until the opening of the Red Sea. In contrast, the sub-African mantle, which had not been extensively depleted by such late stage calc-alkaline magmatism retained its low melting point components and continued to produce small quantities of alkaline magma throughout the Phanerozoic. During the Tertiary the Red Sea rift developed along the line separating depleted and non-depleted upper mantle, which in turn may coincide with a Proterozoic suture zone.

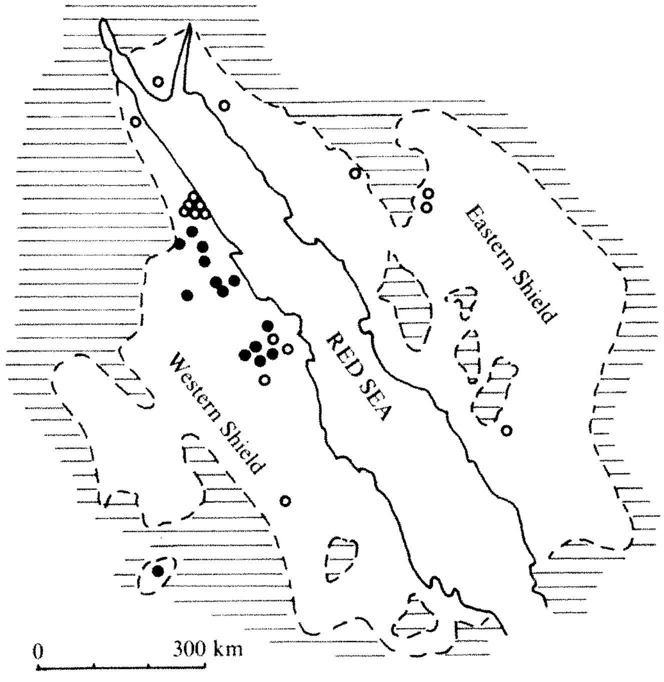


Fig. 2 Distribution of dated samples from alkaline and peralkaline intrusions. ○, 660–540 Myr. ●, 480–80 Myr. Horizontal shading marks Phanerozic cover.

Received 2 June; accepted 28 November 1980.

1. Gass, I. G. *Nature* **165,** 722–724 (1977).
2. Steigler, R. H. & Jager, E. *Earth planet. Sci. Lett.* **36,** 359–363 (1977).
3. Fleck, R. J., Greenwood, W. R., Hadley, D. G., Anderson, R. E. & Schmidt, D. G. *U.S. geol. Surv. Saudi Arabian Proj. Rep.* No. 245 (1979).
4. Aldrich, L. T., Brown, G. F., Hedge, C. & Marvin, R. F. *U.S. geol. Surv. Proj. Rep.* No. 240 (1978).
5. Baubron, J. C., Delfour, J. & Vialette, Y. *Open-file Rep. 76-JED-22* (Bureau de Recherches Geol. Miniers, 1976).
6. Bielski, M., Jager, E. & Steinitz, G. *Contr. Miner. Petrol.* **70,** 159–165.
7. Roger, J. J. W., Ghuma, M. A., Nagy, R. M., Greenberg, J. K. & Fullager, P. D. *Earth planet. Sci Lett.* **39,** 109–117 (1978).
8. Lutz, T. M., Foland, K. A., Faul, H., Friedman, I. & Serencsits, C. McC. *U.S. geol. Surv. Open-file Rep. 78–701*, 465–467 (1978).
9. Cavanagh, B. J. thesis, Univ. Leeds (1979).
10. Hashad, A. H. *Bull. Inst. appl. Geol. Jeddah*, **3,** 31–46 (1980).
11. Vail, J. R. *Geol. J.* **B20,** 97–114 (1976).
12. Gass, I. G. thesis, Univ. Leeds (1955).
13. Serencsits, C. McC, El Ramly, M. F., Faul, H., Foland, K. A. & Hussein, A. A. *Annls. geol. Surv. Egypt* (in the press).
14. El Ramly, M. F., Budanov, V. I., Armanious, L. K. & Devenuik, N. E. *United Arab Repub. Geol. Surv.* Pap. 52 (1969).
15. Stoeser, D. B. & Elliot, J. E. *U.S. geol. Surv. Proj. Rep.* No. 265 (1979).
16. Harris, N. B. W. & Marriner, G. F. *Lithos* **13,** 325–337 (1980).
17. Bailey, D. K. in *The Alkaline Rocks* (ed. Sovensen, H.) (Wiley, New York, 1974).
18. Moore, J. M. *J. geol. Soc. Lond.* **136,** 441–454 (1979).
19. Wyllic, P. J. & Huang, W. L. *Contr. Miner. Petrol.* **54,** 79–107 (1976).
20. Gass, I. G. *J. geol. Soc. Lond.* **136,** 129–138 (1977).
21. Greenwood, W. R., Hadley, D. G., Anderson, R. E., Fleck, R. J. & Schmidt, D. L. *Phil. Trans. R. Soc.* A**280,** 517–526 (1976).

Solar energy conversion by chloroplast photoelectrochemical cells

Ravindra Bhardwaj, Rong L. Pan & Elizabeth L. Gross

Department of Biochemistry, The Ohio State University,
484 West 12th Avenue, Columbus, Ohio 43210

Solar energy can be captured by photochemical or photoelectric processes in which a gradient of chemical potential can be generated using photochemical redox reactions. Photosynthesis is the most efficient system for quantum conversion and storage of solar energy[1]. Various biomimetic systems have been constructed as models for the conversion of light to electrical energy[2–11]. These biological photoelectrochemical cells have utilized microcrystalline chlorophyll *a* (Chl *a*)[2], chlorophyll liquid crystals deposited on the electrodes[3–5] or incorporated into filters[6], bilayer lipid membranes[7], chloroplast membranes[8], chloroplasts[9] and bacterial reaction centres[10,11]. The power conversion efficiencies are mostly very low (≈0.002%). We describe here a chloroplast photoelectrochemical cell which generates large photovoltage and photocurrent, and has a power conversion efficiency close to 1%.

We have previously developed a solar cell using photosystem I (PSI) particles[12]. Our rationale was that charge separation occurs within the reaction centre of PSI particles and produces a

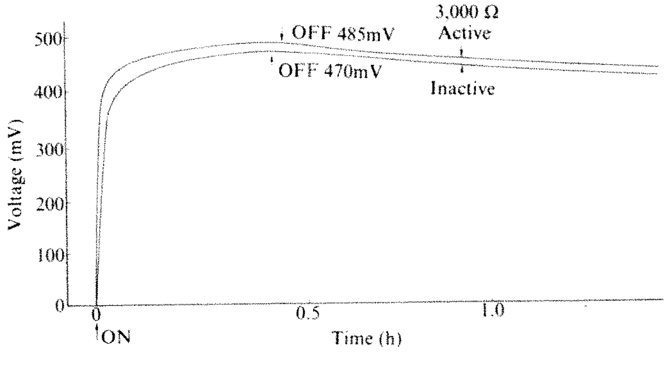


Fig. 1 Time course of potential development. Chloroplasts containing 80 μg Chl were aspirated onto a 1-cm^2 area (1.2-μm pore diameter) Metrical filter. The cell contained 80 mM Tricine (*p*H 7.0) + 100 mM NaCl + either 3.0 mM FMN or 0.5 mM DCIP. The FMN side (front surface area = 8 cm^2) was illuminated with white light ($I = 4,600$ W m^{-2}). The potential was measured across a 3,000 Ω load resistance.

0028-0836/81/050396—03$01.00

Table 1 The effect of buffers on the chloroplast solar cell

Buffer	Buffer concentration (mM)	Condition	DCIP/chloroplasts/FMN *p*H 7.0				DCIP/chloroplasts/AQS *p*H 8.2			
			Maximum power (μW)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)	Maximum power (μ W)	Resistance at maximum (power (Ω)	V_{oc} (mV)	I_{sc} (mA)
Tris-Cl	20	(+)	149	200	390	2.2	134	500	442	1.5
		(−)	117	200	380	1.9	103	500	435	0.95
	80	(+)	192	200	405	2.6	242	200	452	2.3
		(−)	151	200	395	2.0	143	600	443	1.2
Na-Tricine	20	(+)	345	200	480	3.2	159	500	452	1.6
		(−)	276	200	475	2.5	103	700	445	1.1
	80	(+)	384	100	515	3.6	232	300	458	2.2
		(−)	294	200	510	2.8	140	600	452	1.3
Phosphate	20	(+)	151	200	430	2.0	22	6,000	465	0.30
		(−)	111	300	430	1.7	2.0	900	360	0.19
	80	(+)	158	200	430	2.15	32	4,000	475	0.35
		(−)	119	300	430	1.75	2.5	900	380	0.20

FMN (3.0 mM) and AQS (1.5 mM) were used as acceptors. DCIP (0.5 mM) was the donor. (+) and (−) indicate active and inactive (killed) chloroplasts respectively.

strong reductant (reduced P430)[13] and a weak oxidant (P700$^+$)^{5}g3[14]. The reductant and oxidant then interact with exogenous electron acceptors and donors producing a gradient of chemical potential which can be converted into an electrical current. Also, the electron acceptor used (flavin mononucleotide, FMN) is photoreducible either directly by the electron donor (Fe^{2+}) in the case of potassium ferrocyanide[15] or by Tricine[16]. Thus, both photosynthetic reduction of FMN mediated by PSI and FMN photochemistry (FMN photoreduction) contribute to the power developed. In the case of FMN, photodestruction also occurs. Using isolated chloroplasts in place of the PSI particles in the photochemical cell has proved advantageous because of their ease of preparation and their power conversion efficiency of close to 1%.

Isolated PSI particles or broken chloroplasts (Type C) were placed on a filter between two compartments containing the electron donor and the electron acceptor, respectively. Platinum electrodes were placed in each compartment and the photovoltage was observed across a load resistance (see ref. 12 for experimental details and design). Initially[12], we obtained an open circuit photovoltage (V_{oc}) of 360 mV, a short circuit current (I_{sc}) of 108 μA and a maximum power (W_{max}) of 20 μW.

We determined the characteristics of our chloroplast photoelectrochemical cell in the presence of two photoactive electron acceptors, FMN and AQS (anthroquinone 2-sulphonate) in optimal conditions of donor (dichlorophenol indophenol, DCIP) and acceptor concentrations, *p*H and buffer strength (Table 1). In all cases, the chloroplasts contribute at least 30% of the power observed. The two acceptors have different *p*H optima: 7.0 for FMN and 8.2 for AQS. Both Tris and Tricine buffers increased the power observed because they act as electron donors for the photoreduction of FMN[16] and AQS[17]. The increase in power with increasing buffer concentration is not a salt effect because increasing NaCl concentration from 100 to 200 mM had no effect. The AQS system did not generate greater power in phosphate buffer because, unlike FMN, there is no photodestruction which also contributes to the power developed. FMN was a better electron acceptor than AQS because of its greater absorption cross-section in the visible region of the spectrum and the fact that photodestruction could contribute to the potential developed. A maximum power of 400 μW was observed across a 100 Ω resistance. The open circuit voltage and short circuit current were 510 mV and 3.5 mA respectively.

The time course of the photoresponse is shown in Fig. 1. A photovoltage of 485 mV was observed across a 3,000 Ω resistance. That the potential persisted for 1–2 h even after the light was turned off, illustrated the system's ability to store energy. The ability of the cell to generate a voltage is equivalent to a generator in series with a source resistance (R_s) whose magnitude is a reflection of the ability of the cell to generate current. The R_s in a photoelectrochemical cell is a function of rate of electron transport through PSI during illumination and the rate of photoreduction of the acceptor (FMN or AQS) by the reducing agents. The more efficient these processes, the lower R_s will be. Because both photosynthetic and photochemical reduction of the acceptor depend on light, R_s should very high in dark and low in light. R_s also operates in parallel with membrane resistance (R_m). To develop and store energy, R_m has to be very high to limit the dissipation of the reduced photoproducts through the membrane. We have measured the R_m and the differences in the internal resistance in the dark and during steady state illumination (Fig. 2). The R_m was calculated to be of the order of 1.5×10^5 Ω (Fig. 2, curve 1). In the dark, the internal resistance was 5×10^4 Ω (curve 2) which decreases to 150 Ω (curve 3) after steady state photovoltage is achieved during illumination. The

Table 2 Comparison of chloroplasts and PSI particles

Donor (mM)	Acceptor (mM)	*p*H	Condition	Chloroplasts				PSI			
				Maximum power (μW)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)	Maximum power (μW)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)
DCIP (0.5)	FMN (3.0)	7.0	(+)	384	100	515	4.0	432	100	521	4.4
			(−)	294	200	510	3.0	342	100	510	3.3
DCIP (0.5)	AQS (1.5)	8.2	(+)	232	300	458	2.1	306	100	460	3.3
			(−)	140	600	452	1.5	192	300	460	2.4

Chloroplasts (containing 80 μg Chl were aspirated onto 1-cm^2 filter of 1.2 μm pore diameter. Conditions included 80 mM Na-tricine (at the indicated *p*H). The light intensity at the front of the cell was 5,000 W m^{-2}.

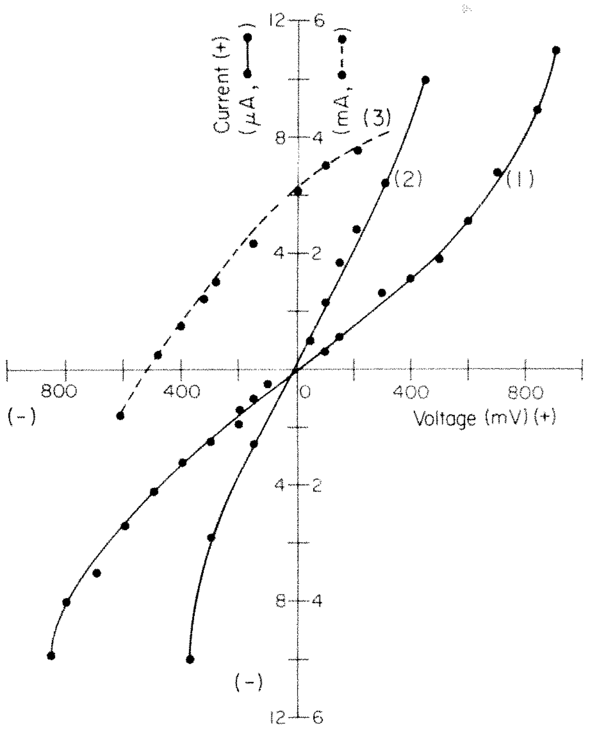


Fig. 2 Current–voltage curves of PSI impregnated membrane in absence of any dye (curve 1), in presence of dyes in dark (curve 2) and during illumination in presence of dyes after steady state potential was developed. Controlled potential was applied using a battery and a potentiometer and the current was measured. Membrane resistance (R_m) (curve 1) was measured only in salt solution. Other conditions as in Fig. 1. Light intensity was 2,000 W m^{-2}.

current–voltage curve shows that in light (curve 3) the V_{oc} is 520 mV and I_{sc} is 3.1 mA.

Table 2 shows that the chloroplasts can effectively substitute for the PSI particles. The power and short circuit current obtained is marginally higher in the case of PSI particles. This may be due to a permeability barrier in chloroplasts.

We calculated the power conversion efficiency for the FMN system assuming a total front surface area of 8 cm^2. The efficiency increased as the incident light intensity was lowered from 0.08% at 5,000 W m^{-2} to 0.4% for 30 W m^{-2}. The contribution of the chloroplasts to the total power also increased as the light intensity was lowered from 30% at 5,000 W m^{-2} to 50% at 30 W m^{-2}. If we calculate the efficiency for 30 W m^{-2} based on the 1 cm^2 area of chloroplasts illuminated rather than the 8 cm^2 of FMN solution, we obtained an efficiency of 3.2%. The true efficiency lies between the two estimates. By subtracting the photochemical contribution from the total power an efficiency of 2.3% was obtained for the chloroplast contribution. This may not be correct, however, as the two contributions are synergistic.

This work was supported by the US Department of Energy, grant DE FG02-79ER10538. The views presented here do not necessarily reflect those of the DOE.

Received 4 August; accepted 20 November 1980.

1. Bolton, J. R. in *Proc. 4th int. Congr. on Photosynthesis* (eds. Hall, D. O. Coombs, J. & Goodwin, T. W.) 621–634 (The Biochemical Society, London, 1978).
2. Tang, C. W. & Albrecht, A. C. *Nature* **254,** 507–509 (1975).
3. Aizawa, M. *et al. J. Membrane Sci.* **4,** 251–259 (1978).
4. Aizawa, M. *et al. Electrochim. Acta* **24,** 89–94 (1979).
5. Mountz, J. M. & Tien, H. T. *Solar Energy* **21,** 291–295 (1978).
6. Tien, H. T. & Mountz, J. M. *Energy Res.* **2,** 197–200 (1978).
7. Tien, H. T. *Photochem. Photobiol.* **24,** 97–116 (1976).
8. Allen, M. J. & Crane, A. E. *Bioelectrochem. Bioenerg.* **3,** 84–91 (1976).
9. Haehnel, W. & Hochheimer, H. J. *Bioelectrochem. Bioenerg.* **6,** 563–574 (1979).
10. Drachev, L. A. *et al. FEBS Lett.* **50,** 219–222 (1975).
11. Janzen, A. F. & Seibert, M. *Nature* **286,** 584–585 (1980).
12. Gross, E. L. *et al. Photochem. Photobiol.* **28,** 249–256 (1978).
13. Ke, B. *Biochim. biophys. Acta* **301,** 1–33 (1973).
14. Norris, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **68,** 625–628 (1971).
15. Rutter, W. J. *Acta Chim. scand.* **12,** 438–446 (1958).
16. Nelson, N. *et al. Photochem. Photobiol.* **16,** 481–489 (1972).
17. Robinson, H. H. & Yocum, C. F. *Photochem. Photobiol.* **29,** 135–140 (1979).

Liquefaction of biomass as a source of fuels or chemicals

Ian J. Miller & S. Kevin Fellows

Chemistry Division, Department of Scientific and Industrial Research, Petone, New Zealand

Although biomass is one of the few renewable sources of liquid fuels and chemicals, relatively little work has been carried out on its direct liquefaction. Here, we report that wood or cellulose can be almost totally converted to liquids or gases when heated under pressure to 350 °C in the presence of phenol, water, an acid or base catalyst, and either hydrogen or nitrogen. The resultant liquids were separated into a neutral fraction, which was partly characterized by gas chromatography–mass spectrometry (GC–MS) and found to contain mainly aromatic compounds with or without an oxygen-containing ring, and a phenolic fraction. The nature of these compounds was found to be dependent on the catalyst, but not appreciably on the physical conditions used.

Previous attempts at liquefying biomass have tended to consider it as a substitute for coal. Thus liquids have been obtained by pyrolysis under hydrogen[1], by pyrolysis in wood tars, by reaction with carbon monoxide[2], by hydrogenation with iodine[3] or nickel[4] as a catalyst, or by hydrogenation with a 'donor solvent'[5]. These processes, however, fail to take advantage of the depolymerization of carbohydrate fraction at least at ~350 °C to monomer units such as 1,6-anhydroglucose[6]. This is followed by degradation and condensation reactions which, if unchecked, lead to the formation of char. These intermediates need therefore to react or to be rearranged before they can condense.

A preliminary objective is to slow down the condensation reactions, which is best done by providing a solvent to slow down the higher order solid state reactions. We used phenol because as the hydrogenation of lignins produces mainly phenols[7] they would be a significant by-product of the process, and phenol with certain catalysts such as zinc chloride, dissolves cellulose. To interrupt condensation reactions we first tried hydrogenation, and then simple acid–base catalysed rearrangements.

In a typical qualitative reaction 2 g of the cellulosic material, 2 g of phenol, 2.5 g of water, and 0.5 g of other reagents were placed in a glass liner which in turn was placed in a 15 cm^3 capacity pressure vessel. This was pressurized to 3.4 MPa with either hydrogen or nitrogen, and heated to 350 °C within 30 min for times between a few minutes and several hours. After cooling, the liquid products were separated into phenolic and non-phenolic fractions, and the major components of the latter characterized by GC–MS. Some of the phenolic fraction was also characterized, and was similar overall to that reported from the hydrogenation of lignin[7].

The composition of the neutral fraction did not depend appreciably on the source of the biomass. When heated with zinc chloride and nitrogen, cellulose (filter paper, cut into squares), softwood (*Pinus radiata*, chips ≤6×6×2 mm) or hardwood (aspen, powdered, similarly to Masonite process, by steam explosion) gave products of almost identical composition.

The biomass was required to remain in contact with the solvent. If the volume of solvent was reduced below that outlined above, charring occurred. Only 1 g of wood chips could be used as they did not pack very well. The powdered aspen was used for most experiments because it packed better, and it could be slurried with the two phases of the liquid. To mimic green biomass, aspen with 40% dry solids, 60% moisture was used.

Liquefaction of powdered aspen was essentially complete after a few minutes at 350 °C. The composition of the neutral fraction was unaffected by reaction time, although the yield was

0028-0836/81/050398—02$01.00

Table 1 Liquid products obtained from aspen at 350 °C for 1 h, with phenol, water and other reagents

	% Composition of neutral products																
Reagents	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
$ZnCl_2$, N_2	8	0.6	4	1	2	62	2										
$ZnCl_2$, Ni, N_2	0	0.6	5	1	2	58	10										
$ZnCl_2$, H_2, Ni	15	15	21	1		31		1*	3*	3*	0.5*	0.5*					
NaH_2PO_4, N_2	5	4		0.5		11		7.5†					43	6			
Na_2CO_3, H_2, Ni	5	4		0.5		19		7.5†							5	17	5

Liquid products: (1), anisole; (2), benzofuran; (3), 2-methylbenzofuran; (4), dibenzofuran; (5), diphenylether; (6), xanthene; (7), xanthone; (8), 2,3-dimethylbenzofuran; (9), ethylbenzofuran; (10), 1,2-dimethylindane or methylindanone; (11), methylxanthene; (12), methylxanthenone; (13), tetralin; (14), naphthalene; (15), cyclohexanone; (16), phenylcyclohexen-3-one; (17), ketone, tentatively phenylcyclohexan-3-one.

* Tentative identification.

† Products (8) and (9) combined.

lower at shorter times (with evidence of cellulose pyrolysates present), and at times appreciably longer than an hour, when tarring occurred. Reactions below 325 °C led mainly to char formation, and the time between 250 and 325 °C had to be minimized. This problem could be alleviated by shaking and by having larger amounts of phenol present. Reactions over 400 °C led to tarring.

Representative results using aspen are shown in Table 1. If zinc chloride was used with nitrogen, the main non-phenolic products were as listed in the first row of Table 1. If small amounts of nickel were added, there were increases in the amounts of 2-methyl benzofuran and xanthone, and a decrease of xanthene (Table 1). This may indicate that some hydrogenation occurs even without the addition of hydrogen gas, hydrogen presumably coming from the lignin. If hydrogen were used instead of nitrogen, the same products occur, but in significantly different proportions, and other compounds could be tentatively identified. About 3% of aromatic hydrocarbons, including toluene, ethyl benzene, and xylenes were also formed.

Replacing the zinc chloride with either sodium dihydrogen phosphate or sodium carbonate also gave complete liquefaction. Results and the conditions used are given in Table 1. Copper, cobalt, iron, and chromium were also used in place of nickel, and all gave liquefaction, with varying proportions of the same products.

For quantitative experiments, 2.5 g oven dried aspen, 10 g phenol, 13.5 g water, 0.8 g of a reagent and, if required 0.4 g $Ni(OH)_2$ were slurried, and placed in a glass liner for a larger reactor, which was pressurized to 4 MPa and heated to 350 °C in 40 min. The reagents used, and the recovered yields of neutrals were: zinc chloride/N_2, 0.76 g; zinc chloride/Ni/H_2, 0.74 g; NaH_2PO_4/N_2 0.55 g, and K_2CO_3/N_2, 0.65 g.

Positive yields of phenol were obtained in two reactions, namely $ZnCl_2/H_2/Ni$ (10.2 g recovered), and K_2CO_3/N_2 (10.15 g recovered). The phosphate reacted with the glass, and there was extensive char formation on the exterior of the liner. This was assumed to have come from phenolics reacting with alkaline glass, but in any case is irrelevant to further process development. That the phenol could be recycled was shown by reacting 1-inch squares of filter paper with zinc chloride, nickel, and hydrogen. The phenol was recycled three times before handling losses made it unlikely that the cellulose would dissolve properly. In these conditions 25 g of cellulose gave 12 g of neutral liquids which distilled between 50 and 400 °C.

If cellulose alone is converted to the products in Table 1 by the abstraction of oxygen atoms, the maximum yield is ~60% by weight, whereas if the conversion proceeds by the elimination of carbon dioxide, the maximum yield is ~40%. With a lignin content of 25%, the carbohydrates of aspen would give ~1.1 or 0.7 g of products if each conversion were totally efficient. The maximum yield of phenolics from the lignin would be ~0.7 g, although polyhydroxy phenols would not have been recovered in this study. Handling losses with the original solvent also prevent a proper assessment of phenol yield. The observed recoveries do establish, however, that the reaction is a net producer of phenols.

In conclusion, cellulosic material can be liquefied in conditions which, compared with those required for coal liquefaction, are relatively mild and fast. For a given set of catalysts, ~75% of the non-phenolic products consist of less than eight chemical compounds, and in some cases even fewer. Furthermore, a change in the catalysts can significantly alter the nature of the products. Thus cellulosic feedstocks could be the basis of a significant range of chemical products. Also, because hydrogen is not a necessary component of the liquefaction process for some products, it is possible that crude feedstock for a fuels or chemical industry could be manufactured in many small satellite reactors, thus obviating the large transport costs incurred if an equivalent amount of biomass were processed at one site. The satellite reactors could be low in cost, while the central processing unit could still manage the economies of scale necessary for product costs comparable with materials such as coal.

We thank Steven Gwyn for GC–MS analyses and assistance with product identification.

Received 26 August; accepted 26 November 1980.

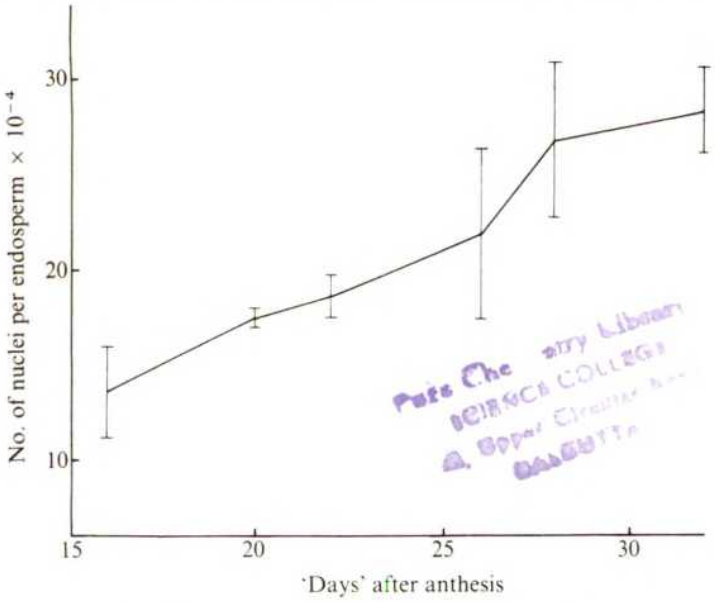
1. Fierz-David, H. E. *Chemy Ind.* **44,** 942–4 (1925).
2. Appel, H. R., Fu, Y. C., Friedman, S., Yavorsky, P. M. & Wender, I. *U.S. Bureau of Mines* RI 7560 (1971).
3. Geerards, J. J. Th. M., van Krevelen, D. W. & Waterman, H. I. *Brennst.-Chem.* **39,** 11–14 (1958).
4. Kaufman, J. A., Gupta, D. V., Szatkowski, T. S. & Wiess, A. H. *Chemie Ingr.-Tech.* **46,** 609 (1974).
5. Boomer, E. H., Argue, G. H. & Edwards, J. *Can. J. Res.* **13B,** 323–350 (1935).
6. Shafizadeh, F. *Adv. Carbohyd. Chem.* **23,** 419–474 (1968).
7. Stamm, A. J. & Harris, E. E. *Chemical Processing of Wood* Ch. 17 (Chemical Publishing Co., New York, 1953).

Endosperm cell number in barley

M. P. Cochrane & C. M. Duffus

Department of Agricultural Biochemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK

Grain weight is one of the major components of grain yield and it has been suggested that in wheat, final grain weight may be a function of endosperm cell number[1]. If grain weight is to be increased by increasing the rate and/or duration of endosperm cell division, it is important to establish the stage of development at which maximum endosperm cell number is attained. Evers[2] reported that no cell division occurred in the endosperm of wheat later than 16 days after anthesis and other workers[3,4] have come to a similar conclusion using cell counts. However, biochemical analyses[5] have shown that the maximum DNA content of wheat endosperm is not reached until 25 days after anthesis and both Donovan[6] and Radley[7] have suggested that endosperm cell division may continue until about 30 days after anthesis. We have estimated the number of endosperm nuclei in developing caryopses of barley and have examined transverse sections of endosperms of different ages. We report here that cell division in barley endosperm continues until 28–30 days after anthesis.



Pate Che stry Library
SCIENCE COLLEGE
92, Upper Circular Rd.
CALCUTTA

Fig. 1 The number of nuclei per endosperm during barley caryopsis development. Bars indicate confidence limits at $P = 0.05$.

In the present investigation, as in previous work in this laboratory[8], barley (*Hordeum vulgare* var. *distichum* L. cv. Midas) was grown in a greenhouse, and caryopsis ages were estimated on a scale based on the normal development of field-grown barley in Scotland, which takes 60 days from anthesis to harvest-ripeness. The parameters used to estimate developmental age include pericarp shape and colour, endosperm shape and colour, and embryo length. The middle three or four grains were removed from each of several of the first-maturing ears in two or more pots. The developmental age of each caryopsis was determined and the endosperm dissected out, taking care to remove the testa and embryo and leave the aleurone intact. It was not possible to sample caryopses older than 33 'days' after anthesis because at this stage of development, the aleurone adheres to the testa and peels off with it. Endosperms were fixed overnight in 1 : 3 v/v acetic acid/ethanol and then stored in 70% ethanol. Endosperms of the same developmental age were pooled and three samples of 10 were stained using Feulgen reagent and digested in cellulase[9]. Cells were suspended in 0.5 M sorbitol and nuclei were counted in a haemocytometer. Caryopses from the same ears were prepared for light microscopy[8].

The number of nuclei per endosperm increased during the period 16–28 'days' after anthesis (Fig. 1). Transverse sections of mid-grains (Fig. 2) show that the aleurone layer of a 15-'day' caryopsis is one to two cells thick, with cell division occurring in both anticlinal and periclinal planes. The cells in the inner and outer layers of the testa are similar in size and the cells of the nucellus are fully expanded and thin-walled. The endosperm of the 15-'day' caryopsis is flat and greyish-white, and the embryo cannot be seen with the naked eye. In 23-'day' caryopses the aleurone is two to three cells thick and cell division can be observed (Fig. 2). The cells in the outer layer of the testa are thinner than those in the inner layer, and the nucellar cells are thick-walled. At this stage the endosperm is white and rounded and the embryo is ~0.25 mm long. By 33 'days' after anthesis, three to five layers of aleurone are established and clearly differentiated from the sub-aleurone and the starchy endosperm. The inner testa cells are large and have osmiophilic contents, and the nucellus is represented only by a layer of compressed wall material. In 33-'day' caryopses the endosperm is yellowish, firm and well rounded, the embryo is ~2.1 mm long and the aleurone layer peels off with the testa.

These observations on developing barley caryopses differ from those reported for wheat in that the number of nuclei per endosperm in barley increased up to about 30 'days' after anthesis, whereas in wheat, increase in the number of nuclei per endosperm stopped 16–20 days after anthesis[3,4]. In addition, evidence of endosperm cell division was observed at a later stage of development in barley than that reported for wheat[2]. These differences may be due to: (1) some assays being done on wheat endosperms contaminated with testa or stripped of their aleurone cells; (2) comparisons being made on the basis of strict chronological age, taking no account of variation in morphological development due to environmental factors; (3) the fact that the aleurone layer in wheat caryopses is only one cell thick whereas that in barley is at least three cells thick and so cell division in barley may continue for some time after the last starchy endosperm cells have been initiated. Note, however, that there is evidence of anticlinal divisions in the meristematic aleurone of 23-'day' barley endosperms and the cells thus produced presumably contribute to an increase in the surface area of the endosperm and hence to an increase in endosperm volume. It seems unlikely that similar anticlinal divisions in the aleurone of wheat cease as early as 16 days after anthesis[2].

Perhaps, therefore, more attention should be paid to the suggestions of Donovan[6] and Radley[7] that cell division in wheat endosperm continues up to 25–30 days after anthesis and to

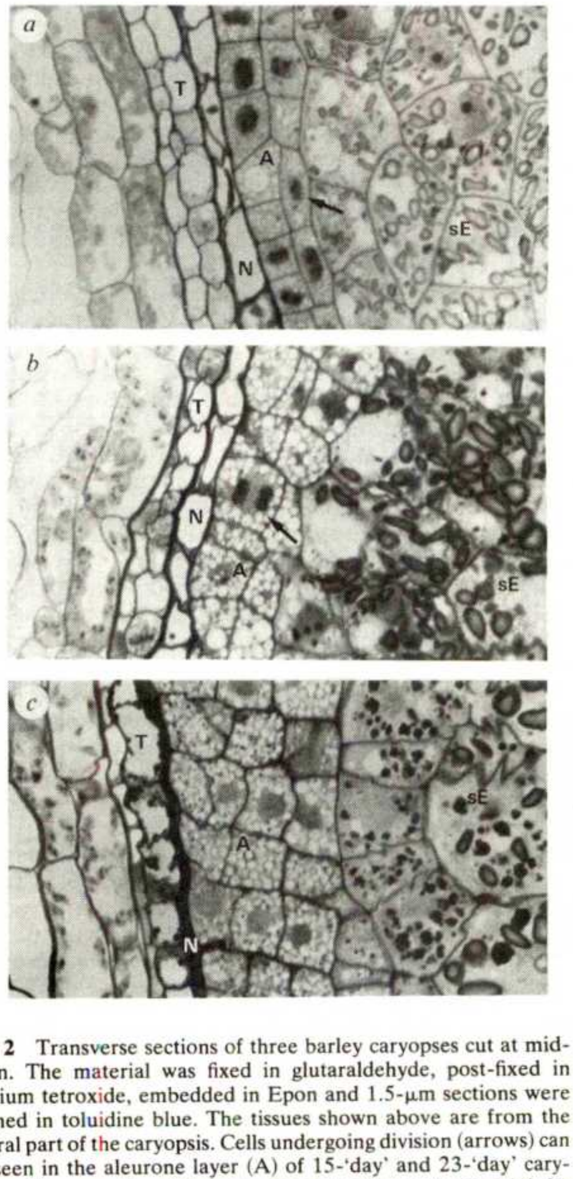


Fig. 2 Transverse sections of three barley caryopses cut at mid-grain. The material was fixed in glutaraldehyde, post-fixed in osmium tetroxide, embedded in Epon and 1.5-μm sections were stained in toluidine blue. The tissues shown above are from the lateral part of the caryopsis. Cells undergoing division (arrows) can be seen in the aleurone layer (A) of 15-'day' and 23-'day' caryopses. These sections illustrate the state of development of the nucellus (N), testa (T) and starchy endosperm (sE) as the number of endosperm nuclei increases (Fig. 1). *a*, 15-'day'; *b*, 23-'day'; *c*, 33-'day'. ×400.

Sandstedt[10], who observed that cell divisions occurred in the endosperm of wheat, grown in Nebraska, up to 16 days after anthesis, that is, half-way between anthesis and maturity in those climatic conditions. If attempts are to be made to increase the weight of cereal grains by applying chemicals which increase endosperm cell number, it is important to establish the pattern of cell division in the developing grain. The timing of the application of any such chemicals could be particularly critical in barley because it seems that cell division taking place after a certain stage of development produces mainly aleurone cells and therefore does not provide further capacity for carbohydrate storage. Thus, stimulation of endosperm cell division early in development could increase the yield of carbohydrate, whereas stimulation of cell division at a later stage could improve the nutritional quality of the grain by increasing the yield of the proteins, fats, minerals and vitamins which are found in aleurone cells.

This work is supported by Agricultural Research Council grant AG/15/169.

Received 14 July; accepted 25 November 1980.

1. Brockelhurst, P. A. *Nature* **266**, 348–349 (1977).
2. Evers, A. D. *Ann. Bot.* **34**, 547–555 (1970).
3. Wardlaw, I. F. *Aust. J. biol. Sci.* **23**, 765–774 (1970).
4. Briarty, L. G., Hughes, C. E. & Evers, A. D. *Ann. Bot.* **44**, 641–658 (1979).
5. Donovan, G. R., Hill, R. D. & Lee, J. W. *Cereal Chem.* **54**, 546–556 (1977).
6. Donovan, G. R. *Aust. J. Pl. Physiol.* **6**, 449–457 (1979).
7. Radley, M. *J. exp. Bot.* **29**, 919–934 (1978).
8. Cochrane, M. P. & Duffus, C. M. *Ann. Bot.* **44**, 67–72 (1979).
9. Rijven, A. H. G. & Wardlaw, I. F. *Expl Cell Res.* **41**, 324–328 (1966).
10. Sandstedt, R. M. *Cereal Chem.* **23**, 337–359 (1946).

Activation of brown adipose tissue thermogenesis by the ventromedial hypothalamus

Martin N. Perkins, Nancy J. Rothwell, Michael J. Stock & Trevor W. Stone

Department of Physiology, St George's Hospital Medical School,
Tooting, London SW17 0RE, UK

It is well established that electrical stimulation of the ventromedial hypothalamus (VMH) causes a reduction in food intake, whereas electrolytic or chemical lesions in this area result in hyperphagia and obesity in the rat. This has led to the suggestion that either the ventromedial nucleus itself, or nerve fibres passing close by, are important in the control of food intake. However, obesity due to VMH lesions occurs in weanling rats in the absence of hyperphagia and can develop in adult rats pair-fed with controls, indicating that destruction of this area also causes an increased metabolic efficiency (that is, a reduced energy expenditure). In normal rats, hyperphagia induced by feeding a highly palatable cafeteria diet is often accompanied by a large increase in heat production (diet-induced thermogenesis) which tends to prevent excessive weight gain and obesity[1]. This diet-induced thermogenesis is due to sympathetic activation of brown adipose tissue (BAT) and we have now investigated the possibility that the VMH is involved in the activation of this process. We found that electrical stimulation of this area produced increased BAT thermogenesis, which suggests that the VMH exerts a dual influence in the regulation of energy balance—an inhibitory effect on energy intake and a stimulatory effect on thermogenesis and energy output.

BAT thermogenesis can be monitored by recording the tissue temperature, as Flaim *et al.*[2] did when stimulating the sympathetic nerves supplying interscapular BAT. This is the largest and most accessible BAT depot and recordings can be made with minimal surgery. We therefore used the same approach to investigate the effects of electrical stimulation of the VMH on the temperature of interscapular BAT. At the same time we recorded core (rectal) and muscle (quadriceps) temperatures.

Male Sprague–Dawley rats (250–290 g) were maintained on a pelleted stock diet (PRD, Christopher Hill Group) and housed in a metabolism room at 24 ± 1°C. They were anaesthetized with urethane (0.12 g per 100 g) and an array of three coaxial stimulating electrodes ~400 μm apart in the lateral plane was stereotaxically located in the VMH (De Groot's[3] coordinates AP 5.4, L 0.6–0.7, V 9.0 mm), the position of all electrodes being later verified histologically. A small incision was made above the scapulae and a thermocouple probe (Comark 1604) was placed between the two lobes of the interscapular BAT and secured by closing the incision with surgical clips. A similar plastic-coated thermocouple was inserted 5 cm into the rectum and a needle thermocouple (0.5 mm o.d.) was inserted into the left quadriceps muscle. Experiments were carried out in a room kept at 24–25 °C and core temperature remained between 35.5 and 37.5 °C in all animals.

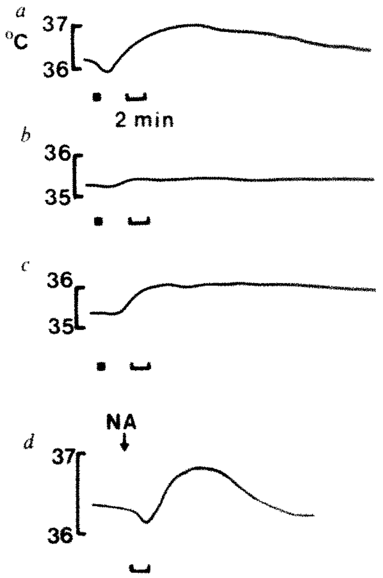


Fig. 1 Records of interscapular BAT temperature following unilateral stimulation of the VMH (■). The electrical stimulus was a 30-s train of monophasic square-wave pulses (0.5 ms, 50 Hz, 60 μA), and histological examination later confirmed that electrodes were located in the lateral and posterior regions of the VMH. *a* Shows a normal response with BAT temperature falling from 36.2 to 36.0 °C 1 min after stimulation, followed by a rise to a peak value of 37.0 °C 9 min later. During this period, rectal and muscle temperatures remained at 36.2 and 35.5 °C, respectively. *b* Shows that treatment with the β-adrenergic antagonist propranolol (0.3 mg per kg, intravenously) reduces the rise in the BAT temperature response to less than 0.2 °C. This effect of propranolol was reversible, as a normal temperature response could be elicited 2 h after treatment (*c*). *d* Shows the temperature response of BAT in a rat given an intravenous injection of noradrenaline (NA, 3 μg per kg). The initial fall in temperature (0.2 °C) and the subsequent rise (0.9 °C above pre-stimulation value) are similar in magnitude and duration to the response to VMH stimulation shown in *a*.

Electrical stimulation of the VMH at hourly intervals produced a biphasic response in interscapular BAT temperature (Fig. 1*a*). The initial decrease in temperature (mean ± s.e.m. = 0.16 ± 0.05 °C, *n* = 5) occurred after 1 min and was followed by a rise (0.86 ± 0.25 °C) which was maximal at 8 ± 2 min and had returned to the pre-stimulation value after 20 ± 4 min. Rectal and skeletal muscle temperatures either did not change or increased slightly (0.25–0.35 °C) after a delay, and on each

0028-0836/81/050401—02$01.00

occasion BAT temperature increased above rectal temperature by 0.3–0.5 °C. The response of BAT to VMH stimulation was remarkably similar to the effects of direct sympathetic nerve stimulation reported by Flaim *et al.*[2]. They also noted an initial small decrease in temperature after 1 min which they ascribed to an α-adrenergic vasoconstrictor effect, and a subsequent large and prolonged increase due to β-adrenergic stimulation of BAT thermogenesis. In our study the α-antagonist phentolamine inhibited the initial decrease in temperature but did not affect the secondary increase produced by VMH stimulation. However, β-adrenergic blockade (using propranolol, Fig. 1*b*) almost abolished the increase in BAT temperature (0.15 ± 0.09 °C, $P < 0.01$ compared with normal response). This confirms the findings of Flaim *et al.*[2] and suggests that VMH activation of BAT thermogenesis also involves β-adrenergic receptors. These findings are consistent with the observation that the increase in metabolic rate of animals exhibiting non-shivering or diet-induced thermogenesis is inhibited by β-receptor blockade but less affected by α-antagonists[1,4].

It is difficult to denervate completely a tissue with a sympathetic supply as rich as that of BAT, and so we used a local anaesthetic to test whether activation by the VMH was mediated by neural or humoral signals. Topical application of tetracaine (1% solution) to the interscapular area abolished the response of BAT to stimulation of the VMH, although the tissue responded to injection of noradrenaline (40 µg per 100 g intraperitoneally) with biphasic temperature responses (Fig. 1*d*) similar in magnitude and duration to stimulation of either sympathetic nerves[2] or the VMH. Central stimulation of thermogenesis by BAT therefore seems to be mediated by the sympathetic nerve supply.

The functional specificity of the VMH has been a matter of debate because, in terms of food intake, the noradrenergic fibres running through this area are apparently more important than the ventromedial nucleus[5]. This may also be true for the effects of VMH stimulation on thermogenesis but, nevertheless, the focus for successful stimulation of BAT thermogenesis can be precisely located in this area. Moving the stimulating electrodes out of the VMH by only 0.5 mm, dorsally or laterally, abolished the response of BAT to central stimulation, and normal responses could be elicited only by moving the electrodes back to their original position. Electrical stimulation in the adjacent lateral hypothalamus had no effect on BAT temperature. Others[6,7] have noted that electrical stimulation of the VMH led to an increase in plasma glycerol and lipogenesis in BAT without affecting lipid synthesis in white adipose tissue, whereas stimulation of the lateral hypothalamus was ineffective. These findings support our suggestion that the VMH is involved in BAT thermogenesis, for diet-induced thermogenesis results in increased BAT lipolysis[1], and lipogenesis in this tissue is considerably increased during exposure to cold[8,9].

The importance of these findings in the regulation of energy balance depends on our proposal[1] that diet-induced thermogenesis has its origins in BAT, and we have demonstrated that the greater thermogenic capacity of hyperphagic rats fed the cafeteria diet[1] is due entirely to increased oxygen utilization by BAT[10]. We have also shown an insulin requirement for diet-induced thermogenesis in hyperphagic rats[11] which, with the results reported here, suggests that changes in glucose utilization in the insulin-sensitive VMH may be responsible for the effects of hyperphagia on diet-induced thermogenesis.

The importance of the VMH in BAT thermogenesis may explain the increased metabolic efficiency (reduced diet-induced thermogenesis) of animals with VMH lesions, particularly in young rats where obesity often develops in the absence of hyperphagia[12]. This explanation is supported by the recent observation that bilateral VMH lesions result in atrophy and reduced thermogenic activity of BAT (L. Girardier, personal communication). It is likely that other areas of the hypothalamus influence BAT thermogenesis, particularly those involved in thermoregulation, but their close proximity to the VMH could provide an explanation for the strong interactions observed between thermogenesis, food intake and environmental temperature[13].

This work was partly supported by the Rank Prize Funds.

Received 6 August; accepted 25 November 1980.

1. Rothwell, N. J. & Stock, M. J. *Nature* **281,** 31 (1979).
2. Flaim, K. E., Horwitz, B. A. & Horowitz, J. M. *Am. J. Physiol.* **232,** R101 (1977).
3. De Groot, T. *Trans. R. Neth. Acad. Sci.* **52,** 1 (1959).
4. Schonbaum, E., Johnson, G. E., Sellers, E. A. & Gill, M. J. *Nature* **210,** 426 (1966).
5. Gold, R. M. *Science* **182,** 488 (1973).
6. Kumon, A., Takahashi, A., Hara, T. & Shimazu, T. *J. Lipid Res.* **17,** 551 (1976).
7. Shimazu, T. & Takahashi, A. *Nature* **284,** 62 (1980).
8. McCormack, J. E. & Denton, R. M. *Biochem. J.* **166,** 627 (1977).
9. Trayhurn, P. *FEBS Lett.* **104,** 13 (1979).
10. Rothwell, N. J. & Stock, M. J. *J. Physiol. Lond.* (in the press).
11. Rothwell, N. J., Stock, M. J. & Warwick, B. *Proc. Nutr. Soc.* (in the press).
12. Bernardis, L. L. & Goldman, J. K. *J. Neurosci. Res.* **2,** 91 (1976).
13. Rothwell, N. J. & Stock, M. J. *Can. J. Physiol. Pharmac.* **58,** 842 (1980).

Photoperiodicity in the male albino laboratory rat

Edward P. Wallen

Department of Life Sciences, University of Wisconsin-Parkside,
Kenosha, Wisconsin 53141

Fred W. Turek

Department of Biological Sciences,
Northwestern University, Evanston, Illinois 60201

Animals inhabiting areas where there are drastic changes in the environment often reproduce only during limited time periods to ensure that young are raised in optimal environmental conditions[1]. The lack of a well defined breeding season in many domesticated animals[2,3], presumably because the selective pressures for seasonal breeding have been minimized, suggests that the neuroendocrine events controlling seasonal cyclicity have been bred out of these animals. Little is known about the underlying neuroendocrine changes which may occur during the evolution of a species from a seasonal to a nonseasonal breeder. Whereas the changing photoperiod is the primary environmental cue which initiates and/or terminates the reproductive season in many animals[4], this is not so in the albino rat *Rattus norvegicus*, a model nonseasonal breeder[5–7]. Nevertheless, daylength can influence various reproductive parameters in laboratory rats [8–17], suggesting that some of the neuroendocrine components that controlled seasonal breeding previously are still extant in this species. To test this hypothesis, we investigated the effect of daylength on the responsiveness of the neuroendocrine–gonadal axis to the negative-feedback effects of testosterone. This paradigm was chosen because of the important role played by photic-induced changes in steroid feedback sensitivity in the control of seasonal reproduction[18–21]. We report here that although daylength has very little effect on neuroendocrine–gonadal function in the intact male laboratory rat, it seems that some component(s) of a photoperiodic system involving the pineal gland has been preserved.

Male inbred Sprague–Dawley albino rats (Gibco, 180–200 g) were housed (four per cage) with food (Purina Rat Chow) and water provided *ad libitum*. Lighting was provided by banks of General Electric cool-white 40-W fluorescent bulbs. The animals were maintained on a light/dark 14:10 h cycle for a 2-week period before starting the experiment. They were then either pinealectomized[22] or sham pinealectomized and implanted subcutaneously with either empty or testosterone-filled Silastic capsules that were 10 or 15 mm long (16 animals per group). Half of the animals in each group remained on LF14:10 while the remaining animals were transferred to LD 6:18. Sixty-three days later the animals were killed by decapitation, blood was collected and the testes removed and weighed. Serum was assayed for luteinizing hormone (LH) and follicle-stimulating hormone (FSH) by methods described previously[23].

0028-0836/81/050402—03$01.00

Consistent with the idea that the male albino laboratory rat is not a photoperiodic species is the observation that testicular weight and serum levels of LH and FSH were similar in untreated animals exposed to either LD 6:18 or 14:10 (Fig. 1, Table 1). On the other hand, in intact rats implanted with 10- or 15-mm long testosterone-filled capsules for 63 days, the testes were significantly ($P < 0.05$) smaller in animals exposed to LD 6:18 compared with those of animals maintained on LD 14:10. The degree of testicular regression observed in the LD 6:18 animals implanted with a 15-mm long testosterone-filled capsule was similar to the maximum regression observed in a previous dose–response study using various sizes of Silastic capsules filled with testosterone[24].

Pinealectomy had no effect on testicular weight in untreated animals exposed to either photoperiod, or in testosterone-treated animals exposed to LD 14:10 (Fig. 1). In contrast, pinealectomy abolished the increased responsiveness to the inhibitory effects of testosterone in animals exposed to short days. Pinealectomized animals maintained on LD 6:18 and implanted with either 10- or 15-mm long testosterone-filled capsules had significantly larger testes ($P < 0.05$) than intact rw on LD 6:18 that were implanted with the same size capsules.

The decrease in testicular weight observed in testosterone-treated rats (Fig. 1) was presumably due to an inhibition of pituitary gonadotropin release[25]. Serum FSH levels were significantly reduced ($P < 0.05$) in all of the testosterone-treated pineal-intact animals when compared with the levels found in animals receiving empty implants. Although serum levels of LH were consistently lower in the pineal-intact rats treated with testosterone than in the untreated control animals, these differences were not statistically significant. However, note that serum LH levels in the intact untreated rats were already near the lower limits of detectability of our assay, and a significant decrease would be difficult to observe.

Surprisingly, no significant effects of either the photoperiod or pinealectomy on serum levels of FSH and LH were observed in testosterone-treated rats (Table 1). This occurred despite the fact that both the photoperiod and pinealectomy did alter the effect of testosterone on reducing testicular size. It is not known how an equal dose of testosterone (that is, 10- or 15-mm testosterone-filled capsules) was able to reduce testis size differentially without also having a differential effect on serum FSH and LH levels. Possibly, a more complete sampling regimen would have yielded differences in gonadotropin levels which were not demonstrable using a single terminal sample. Alternatively, the effect of testosterone may have been via some other hormone which influenced testicular activity in response to the photoperiod and/or pinealectomy. Prolactin is a likely candidate as previous studies indicate that prolactin may be important in the photoperiodic control of testicular activity in rodents[26,27].

Table 1 Measured (mean ±s.e.m.) serum levels of LH and FSH (ng ml^{-1}) in intact and pinealectomized rats

Testosterone capsule size (mm)		LD 14:10		LD 6:18	
		Intact	Px	Intact	Px
0	LH	1.2±0.4	1.2±0.3	1.5±0.4	1.9±0.2
	FSH	336±27	302±17	301±18	345±22
10	LH	0.5±0.2	1.0±0.2	0.6±0.2	0.6±0.1
	FSH	226±20	277±10	220±25	308±37
15	LH	0.6±0.2	0.4±0.1	0.3±0.0	0.9±0.4
	FSH	236±15	289±52	194±15	246±2

The rats were implanted with either empty or testosterone-filled capsules of 10- or 15-mm length. After capsule implantation, the animals were exposed to either LD 14:10 or LD 6:18 for 63 days. LH and FSH values are expressed in terms of nanogram equivalents of NIH-LH-S16 or NIAMMD-Rat-FSH-RP-1.

Px, pinealectomized.

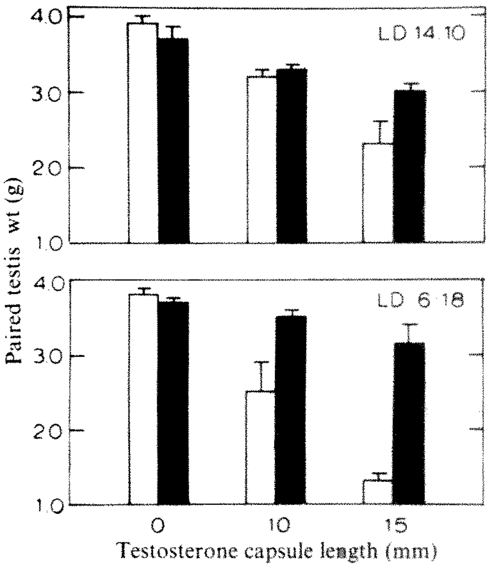


Fig. 1 Mean paired testis weight of rats that were either sham pinealectomized (open bars) or pinealectomized (shaded bars) and implanted with either empty or testosterone-filled Silastic capsules that were 10 or 15 mm long. At the time of capsule implantation the animals were either maintained on LD 14:10 (top panel) or were transferred to LD 6:18 (bottom panel) for 63 days. The vertical line represents the standard error. The testosterone-filled capsules were prepared using Silastic tubing (Dow Corning). Approximately 40 µg of testosterone are released per day per 10 mm of capsule length.

Recent studies on two photoperiodic species, the golden hamster and the sheep, demonstrate the importance of a photic-induced alteration in the responsiveness of the hypothalamic–pituitary axis to the negative-feedback effects of sex steroids in the control of seasonal breeding[18–21,23]—exposure to a nonstimulatory photoperiod renders the animals extremely sensitive to the negative-feedback effects of steroid hormones on pituitary gonadotropin release. Furthermore, in the golden hamster pinealectomy can block the hypersensitivity to testosterone feedback induced by short days[28]. That pinealectomy can block the short day-induced increase in responsiveness to testosterone in both the albino rat and the hamster suggests that the neuroendocrine events associated with photo-induced changes in sensitivity are similar. Indeed, the photoperiod can alter the negative feedback effect of oestrogen on pituitary gonadotropin release in the ovariectomized albino rat[8–10]. However, in the present study, serum levels of LH and FSH were similar in the testosterone-treated pineal-intact rats exposed to LD 14:10 and LD 6:18. Further studies are required to determine how photic-induced changes in sensitivity to steroid hormones in seasonal breeders are related to the change observed in the non-seasonally breeding laboratory rat.

There are conflicting reports regarding the effect of both the photoperiod and pinealectomy on reproductive function in the rat[5–17,29–34]. The present study suggests that such apparent discrepancies may be due to some components of the neuroendocrine system which may have been involved in the photoperiodic control of reproduction at a previous evolutionary time still existing in the rat, and being manifested only in certain experimental conditions. An added complication is that the inbreeding of various strains of laboratory rats may have selected for or against different components of the photoperiodic–neuroendocrine system.

The results reported here on the albino rat are paradoxical. If short days render the animals more sensitive to the inhibitory effects of testosterone, one would expect that testicular regression would occur in untreated rats exposed to short days, as these animals would be more responsive to the testosterone produced by their own testes. One explanation for our results is that the inhibitory effects of short days are blocked or offset by other short day-induced changes which occur in the neuroendocrine system. This hypothesis is particularly attractive because recent evidence demonstrates that at least two mechanisms may be involved in the photoperiodic response of seasonal breeders. In addition to the effects of daylength on steroid-feedback sensitivity, it seems that the photoperiod can also influence pituitary gonadotropin release in the absence of gonadal steroid hormones[35-38]. In rats exposed to short days, there may be neuroendocrine changes which are independent of steroid feedback and which negate the enhanced negative-feedback effects of steroid hormones. Because the photoperiod alters the neuroendocrine–gonadal axis only in certain conditions, the laboratory rat may be useful for the elucidation of the neuroendocrine components involved in the photoperiodic control of reproduction in seasonal breeders.

We thank Sue Losee, John Georgeson and John Smith for technical assistance, Susan Stice for preparing the manuscript, and Gary Ellis for his comments and criticism. This work was supported by NSF grant PCM-7823377 to E.P.W. and NIH grants HD-09885, HD-12622 and Research Career Development Award HD-00249 to F.W.T.

Received 4 August; accepted 21 November 1980.

1. Assenmacher, I. & Farner, D. S. (eds) *Environmental Endocrinology* 1–122 (Springer, Berlin, 1978).
2. Lodge, J. R. & Salisbury, G. W. in *The Testis* Vol. 3 (eds Johnson, A. D., Gomes, W. R. & Vandemark, N. L.) 139–167 (Academic, New York, 1970).
3. Schwartz, N. B. in *The Hypothalamus* (eds Martini, L., Motta, M. & Fraschini, F.) 515–528 (Academic, New York, 1970).
4. Turek, F. W. & Campbell, C. S. *Biol. Reprod.* **20,** 32–50 (1979).
5. Kinson, G. A. & Robinson, S. J. *J. Endocr.* **47,** 391–392 (1970).
6. Reiter, R. J. *et al. Endocrinology* **88,** 895–900 (1971).
7. Kinson, G. A. & Liu, C. *Life Sci.* **14,** 2179–2188 (1974).
8. Hoffmann, J. C. *Biol. Reprod.* **2,** 255–261 (1970).
9. Hoffmann, J. C. & Cullin, A.-M. *Neuroendocrinology* **23,** 285–296 (1977).
10. Wallen, E. P. & Yochim, J. M. *Biol. Reprod.* **11,** 117–124 (1974).
11. McCormack, C. E. *Fedn Proc.* **31,** 811 (1972).
12. Sridaran, R. & McCormack, C. E. *Biol. Reprod.* **20,** 705–712 (1979).
13. Yochim, J. M. & Mitchell, J. A. *Endocrinology* **87,** 465–471 (1970).
14. Mitchell, J. A. & Yochim, J. M. *Endocrinology* **87,** 472–480 (1970).
15. Sorrentino, S. Jr. & Benson, B. *Gen. comp. Endocr.* **15,** 242–246 (1970).
16. Shapiro, M. I., Collu, R., Masse, D., Ducharme, J. R. & Roux, J. F. *Life Sci.* **19,** 1341–1346 (1976).
17. Piacsek, B. E. & Meites, J. *Neuroendocrinology* **2,** 129–137 (1967).
18. Pelletier, J. & Ortavant, R. *Acta endocr.* **78,** 442–450 (1975).
19. Tamarkin, L., Hutchison, J. S. & Goldman, B. D. *Endocrinology* **99,** 1528–1533 (1976).
20. Legan, S. J., Karsch, F. J. & Foster, D. L. *Endocrinology* **101,** 818–824 (1977).
21. Ellis, G. B. & Turek, F. W. *Endocrinology* **104,** 625–630 (1979).
22. Hoffman, R. A. & Reiter, R. J. *Anat. Rec.* **153,** 19–22 (1965).
23. Turek, F. W. *Endocrinology* **101,** 1210–1215 (1977).
24. Berndtson, W. E., Desjardins, C. & Ewing, L. L. *J. Endocr.* **62,** 125–135 (1974).
25. Campbell, C. S. & Schwartz, N. B. *J. Tox. envir. Hlth* **3,** 61–95 (1977).
26. Bex, F. J. & Bartke, A. *Endocrinology* **100,** 1223–1226 (1977).
27. Bartke, A. *et al. Endocrinology* **106,** 167–172 (1980).
28. Turek, F. W. *Endocrinology* **104,** 636–640 (1979).
29. Ronnekleiv, O. K. & McCann, S. M. *Neuroendocrinology* **19,** 97–114 (1975).
30 Alonso, R., Prieto, L. & Mas, M. *Endocrinology* **102,** 1534–1538 (1978).
31. Jacobson, C. D. & Mann, D. R. *Biol. Reprod.* **18,** 712–718 (1978).
32. Brown-Grant, K. & Ostberg, A. J. C. *J. Endocr.* **62,** 45–50 (1974).
33. Blake, C. A. *J. Endocr.* **69,** 67–75 (1976).
34. Takeo, Y., Shirama, K., Shimizu, K. & Maekawa, K. *Annotnes zool. jap.* **49,** 83–89 (1976).
35. Davis, G. J. & Meyer, R. K. *Gen. comp. Endocr.* **20,** 61–68 (1973).
36. Gibson, W. R., Follett, B. K. & Gledhill, B. *J. Endocr.* **64,** 87–101 (1975).
37. Pelletier, J. & Ortavant, R. *Acta endocr.* **78,** 435–441 (1975).
38. Ellis, G. B. *Biol. Reprod.* **20,** Suppl. 1, 30A (1979).

A human quantitative polymorphism related to Xg blood groups

P. N. Goodfellow

Imperial Cancer Research Fund, Lincoln's Inn Fields,
London WC2A 3PX, UK

P. Tippett

MRC Blood Group Unit, Wolfson House, University College,
London NW1 2HE, UK

A monoclonal antibody, 12E7, raised against lymphocytes from a patient with a T-cell acute lymphocytic leukaemia reacts strongly with cortical thymocytes and, to a lesser extent, with other human cells[1]. The antigen defined by 12E7 is not expressed on mouse or hamster cells; this species specificity allowed us to investigate the genetics of the expression of the 12E7 antigen using human–rodent somatic-cell hybrids[2]. Hybrid cells which contain the human X chromosome as the only human genetic contribution react with 12E7. The X-linkage of the gene, or genes, controlling 12E7 antigen expression and the fact that 12E7 antigen is also expressed on red blood cells led us to investigate the possibility of a relationship with the X-linked Xg blood group system. Although the 12E7 and Xg^a antigens seem to be different antigenic determinants specified by different loci, we report here a sex-limited effect involving the Xg phenotype on the expression of 12E7 antigen.

Expression of the Xg blood group system is controlled by two X-borne alleles, Xg^a and a silent allele *Xg*. The antigenic product, Xg^a, of the dominant Xg^a allele is recognized by alloantisera using the indirect antiglobulin haemagglutination method[3,4]. Experiments to demonstrate Xg^a antigen on tissues other than red blood cells have given conflicting results[5,6]. A quantitative polymorphism of the 12E7 antigen on red blood cells was first detected in a family segregating at the *Xg* locus.

The mother J. B., the father W. B. and daughter H. B. all express the Xg^a antigen—this phenotype is designated Xg(a+). The son M. B. has the alternative Xg(a−) phenotype which lacks the Xg^a antigen, indicating that the mother J. B. is heterozygous Xg^aXg. The Xg(a−) son showed a reduced expression of the 12E7 antigen compared with other family members (Fig. 1). The same quantitative polymorphism of 12E7 antigen expression was also found when red blood cells from random donors were tested (Table 1).

Table 1 Results of tests for 12E7 antigen phenotype in 10 random donors

Donor	c.p.m. × 10^{-3} 12E7	c.p.m. × 10^{-3} W6/34	c.p.m. × 10^{-3} CA2.06	Antiglobulin test score 12E7	12E7 phenotype
1	24.0	31.9	2.0	30	High
2	9.1	32.6	1.1	13	Low
3	28.5	33.5	1.2	30	High
4	23.6	37.2	1.1	30	High
5	37.8	37.1	2.2	28	High
6	35.8	42.5	1.4	28	High
7	9.2	33.8	1.1	13	Low
8	34.1	37.4	1.1	30	High
9	36.6	34.4	1.8	30	High
10	8.6	37.6	1.3	0	Low

The binding assay was carried out as described in Fig. 1 legend. 12E7 was used at a dilution of 1 : 200. W6/34, a monoclonal antibody which reacts with a human red-cell antigen, at 1 : 200 dilution, was the positive control. CA2.06, the negative control, was used as described in Fig. 1 legend. The low phenotype was scored if the counts bound with 12E7 were less than half those bound with W6/34 in the same conditions. Repeated testing of several individuals gave consistent results. For phenotypic scoring of 12E7 by the antiglobulin haemagglutination method, see Table 2 legend.

Although the 12E7 antigen was originally detected by the indirect radioimmunoassay (RIA), it could also be demonstrated by an indirect antiglobulin haemagglutination test with the same pattern of high and reduced expression (Table 1); over 50 people were typed by both techniques and complete concordance between the different methods was achieved. Subsequent tests were carried out using the haemagglutination method.

The relationship between the *Xg* locus and the 12E7 polymorphism was investigated by testing over 300 Europeans for Xg^a and for the level of expression of 12E7 antigen (Table 2). All Xg(a+) individuals show high levels of 12E7 antigen expression but Xg(a−) individuals fall into two categories, showing either high- or low-level 12E7 phenotype. The probability that these results occurred by chance is less than 1 in 10^8, indicating a relationship between the *Xg* locus and the 12E7 polymorphism. The simplest explanation for these results is that the 12E7

0028-0836/81/050404—02$01.00

polymorphism is controlled by a previously undetected variant of the silent *Xg* allele. Thus, the Xg^a allele and one *Xg* allele would lead to a high-level expression of 12E7; a different *Xg* allele would give low-level expression of the 12E7 antigen. Alternatively another polymorphic locus might exist with an allele for low-level expression of the 12E7 antigen in very strong linkage disequilibrium with the *Xg* allele. However, both these explanations fail to explain the sex distribution shown in Table 2*b*. Although Xg(a−) males may be high- or low-level expressors of 12E7, Xg(a−) females always have the low-level 12E7 phenotype ($P < 0.001$).

These data imply sex limitation of 12E7 expression. One hypothesis is that there is a polymorphic 12E7 modifying gene and that the allele for high-level 12E7 expression cannot be expressed in females lacking the Xg^a antigen.

An alternative, and perhaps controversial, hypothesis is that a locus, *Yg*, equivalent to *Xg*, is present on the Y chromosome. Like *Xg*, the *Yg* locus must have two alleles, Yg^a and *Yg*. It is postulated that Xg^a and Yg^a allow high-level expression of the 12E7 antigen, whereas *Xg* and *Yg* do not. Thus all Xg(a+) individuals (Xg^a Xg^a, Xg^a *Xg* females and Xg^a Yg^a, Xg^a *Yg* males) and Xg(a−) Yg(a+) males (*Xg* Yg^a) express high levels of 12E7, whereas Xg(a−) females (*Xg Xg*) and Xg(a−) Yg(a−) males (*Xg Yg*) express a low level of 12E7. It is interesting that the gene frequency of the postulated Yg^a allele (0.68), as calculated from Xg(a−) males who are 12E7 high expressors (Table 2*b*), is very similar to that of the Xg^a allele (0.66) in similar populations[5]. We are now searching for the rare families which can be used to distinguish between these hypotheses and are attempting to investigate the biochemical relationship between the Xg^a and 12E7 antigens.

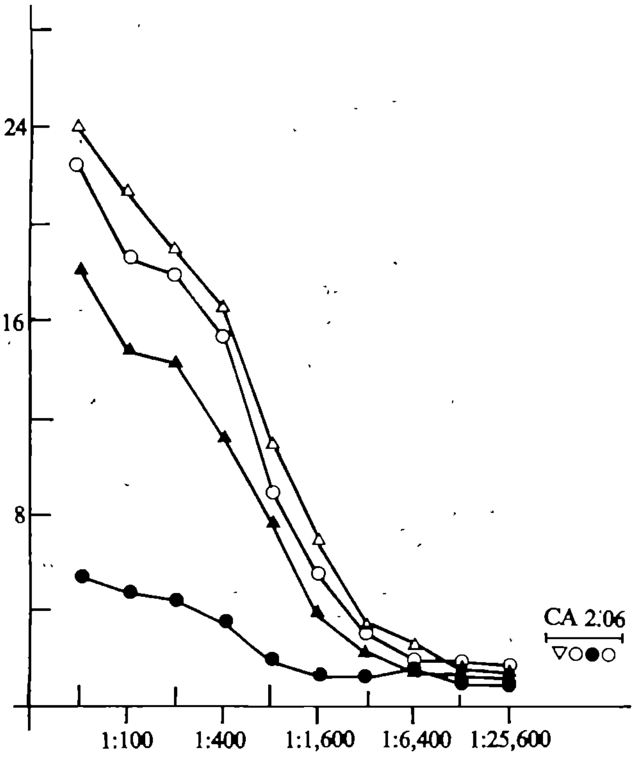


Fig. 1 Binding of 12E7 to red blood cells from a family segregating at the *Xg* locus. 12E7 tests on red cells of: △, father (W.B.); ○, daughter (H.B.); ▲, mother (J.B.); ●, son (M.B.). The indirect RIA was as described elsewhere[2]. Red blood cells carried out (5×10^6) were mixed with 50µl of antibody and incubated overnight at 4 °C The cells were washed and 2×10^5 c.p.m. of ^{125}I-labelled, column-purified, rabbit anti-mouse immunoglobulin added. The mixture was incubated at 4 °C for 1 h. The cells were washed and the bound second antibody quantified in a gamma counter. All dilutions and washes were performed with Dulbecco's phosphate-buffered saline supplemented with 1% bovine serum albumin and 0.1% sodium azide. The 12E7 antibody was prepared by ammonium sulphate precipitation from the cultured hybridoma supernatant[2]. The negative control, CA2.06, a monoclonal anti-human DRw, was used as undiluted culture supernatant[7].

Table 2 Two by two tables comparing Xg^a and 12E7 expression

a		Xg(a+)	Xg(a−)
12E7 expression	High	237	56
	Low	0	42

b		Xg(a−)	
		♂	♀
12E7 expression	High	56	0
	Low	26	16

Phenotypic scoring of 12E7 was carried out 'blind' using either the indirect antiglobulin technique or both indirect RIA and the antiglobulin method. For the latter, washed red blood cells sensitized with 12E7 antibody were tested with rabbit anti-mouse globulin by the capillary method and agglutination was scored at 10, 20 and 30 min. High expression: 25–30; low expression: 0–18. Male Xg(a−) with low expression generally scored more, 0–18, than did female Xg(a−), 0–5.

We thank our colleagues who gave blood for this study; Dr M. Contreras and Mr P. Bowell for providing many blood samples and Drs P. W. Andrews, J. Bodmer, W. F. Bodmer, B. Knowles, R. Sanger, E. Solomon and J. Trowsdale for discussion and guidance. The 12E7 monoclonal antibody was produced and generously donated by Dr R. Levy.

Received 3 September; accepted 18 November 1980.

1 Levy, R, Dilley, J, Fox, R I & Warnke, R *Proc natn Acad. Sci USA* **76**, 6552–6556 (1979)
2 Goodfellow, P., Banting, G, Levy, R, Povey, S & McMichael, A *Somat Cell Genet* **6**, 777–787 (1980)
3 Mann, J D *et al Lancet* **i**, 8–10 (1962).
4 Race, R R & Sanger, R *Blood Groups in Man* 6th edn (Blackwell Scientific, Oxford, 1975).
5 Fellous, M, Bengtsson, B., Finnegan, D. & Bodmer, W F *Ann hum Genet* **37**, 421–430 (1974)
6 Ham, H S, Migeon, B R & Bias, W. B *Human Gene Mapping* 3, *Birth Defects Ser* 12, Vol 7, 382–385 (1976)
7 Charron, D. J & McDevitt, H O *Proc natn Acad Sci USA* **76**, 6567–6571 (1979)

Genetic basis of BCG-induced suppression of delayed hypersensitivity

Denis J. Schrier, John L. Sternick, Elizabeth M. Allen & Vernon L. Moore

Research Service, Wood Veterans' Administration Medical Center, and the Departments of Medicine and Pathology, The Medical College of Wisconsin, Milwaukee, Wisconsin 53193

BCG can either act as an adjuvant to potentiate immunological responses or, in some cases, can induce suppression. The reasons for these differential activities are not clear but may include routes and doses of administration, as well as variable host reactivity to the agent. In this study, we have used killed BCG administered intravenously to produce chronic granulomatous inflammation (CGI) in the lungs and spleen of inbred mice. We report that strains which developed CGI were usually anergic, as evaluated by the development of delayed hypersensitivity (DH) to sheep erythrocytes (SRBC). Studies on the genetics of BCG-induced anergy indicated that it was unigenic, recessive and linked (approximately 28 recombination units) to the immunoglobulin heavy-chain allotype (*Igh*). There was no influence by genes linked to the major histocompatibility complex. The study indicates that anergy associated with CGI is under genetic control, which may explain the variability of anergy in patients with granulomatous diseases. The implication of linkage to the *Igh* complex is not clear, but it may be associated with V_H receptors on T lymphocytes, which in turn act on macrophages to mediate suppression.

We are interested in BCG because it can be used to induce CGI in the lungs of rabbits[1] and in the lungs and spleens of only certain strains of inbred mice[2]. A seemingly paradoxical observation has been that inbred mice which develop CGI in response to BCG are suppressed in their response to mitogens and in the development of DH and a primary antibody response to SRBC[3]. Mice which did not develop CGI were not suppressed in any of these reactivities. Suppression was mediated by cells from inflamed spleens which were adherent to plastic petri plates, lacked Thy 1 and immunoglobulin surface markers, and were removed with carbonyl iron[4]. These traits are characteristic of macrophages.

We examine here the genetic basis of BCG-induced suppression in various inbred mice, in their F_1, F_2 and backcross progeny, and in selected congeneic strains. Data in Table 1 indicate that whereas C57BL/6 (B6) mice (which develop CGI in response to BCG) are markedly suppressed in their ability to develop DH to SRBC, CBA mice (which do not develop CGI to BCG)[2] express DH to SRBC. This is not due to an innate inability of B6 mice to develop DH to SRBC as both strains respond when not treated with BCG (Table 2). Table 1 also indicates that BCG-induced suppression is recessive because BCG-treated (B6 × CBA)F_1 mice were not suppressed. Analysis of F_2 progeny of (B6 × CBA)F_1 mice showed that 26.6% (33/124) were suppressed. Assuming independent segregation, these data suggest that a single gene or gene complex is responsible for suppression. Note also that several of the F_2 progeny of (B6 × CBA)F_1 mice which developed BCG-induced anergy displayed only minimal inflammation. On the other hand, many of the F_2 progeny which were not anergic developed intense inflammation. Although this suggests that genes controlling the development of CGI are different from those which influence anergy, we realise that this is difficult to prove in this system because of the multigenic control of CGI, as contrasted with single gene control of anergy[5]. Because F_1 offspring of either sex from B6 females are not suppressed (data not shown), the control of suppression does not seem to be sex-linked.

To test whether suppression was linked to the *H–2* complex, we performed linkage analysis in F_2 progeny of (B6 × CBA)F_1 parents. The F_2 progeny were typed for either the *H–2*^{b} or *H–2*^{k} allele using anti-*H–2* serum prepared in our laboratory. Anti-*H–2*^{k} serum was produced by injecting B10.BR cells into C57BL/10 mice and anti-*H–2*^{b} serum was produced by injecting C3H.SW cells into C3H mice. The following results were obtained: *H–2*$^{b/b}$, suppressed (10); *H–2*$^{b/b}$ nonsuppressed (18); *H–2*$^{b/k}$ suppressed (11); *H–2*$^{b/k}$, nonsuppressed (40); *H–2*$^{k/k}$, suppressed (4); and *H–2*$^{k/k}$, nonsuppressed (18). These results do not differ significantly from a 6.25 : 18.75 : 12.5 : 37.5 : 6.25 : 18.75 ratio based on independent segregation of a recessive trait ($P > 0.10$). A strain distribution analysis (Table 2) also suggested that genes linked to the *H–2* complex are not involved in suppression. In particular, C57BL/6 (*H–2*^{b}), C57BL/10 (*H–2*^{b}), B10.BR (*H–2*^{k}), and BALB/c.*Igh*^{b} (*H–2*^{d}) all exhibited suppression even though they were different *H–2* types. Table 2 shows also that all strains which were suppressed were *Igh*^{b}, suggesting linkage to the *Igh* allotype. This is further substantiated because BALB/c mice (*Igh*^{a}) were not suppressed, but the congeneic partner, BALB/c.*Igh*^{b}, was. We analysed linkage to the *Igh* complex further by performing linkage analysis on male and female (C57BL/6 × CBA)F_1 × C57BL/6 backcross mice. The backcrosses were typed for either the *Igh*^{b} or *Igh*^{a} allele with the following results: *Igh*$^{b/a}$, nonsuppressed (14); *Igh*$^{b/b}$, suppressed (9); *Igh*$^{b/a}$, suppressed (6); *Igh*$^{b/b}$, nonsuppressed (3). These results differ significantly from a 1 : 1 : 1 : 1 ratio expected on the basis of independent segregation ($P < 0.025$). The estimated recombination frequency is 0.28 ± 0.08.

Table 1 BCG-induced suppression of delayed hypersensitivity to sheep erythrocytes

	Footpad increment (mm)*		
Mouse strain	Positive	Negative	No. suppressed total (%)
C57BL/6		0.16 ± 0.02	8/8 (100)
CBA	0.77 ± 0.06		0/10 (0)
[B6 × CBA]F_1	0.79 ± 0.04		0/22 (0)
[(B6/CBA) × (B6/CBA)]F_2	0.61 ± 0.02	0.17 ± 0.01	33/124 (26.6)†

Mice were injected intravenously with 300 μg of killed BCG in an oil-in-saline emulsion. Four weeks later, B6, but not CBA, mice exhibited intense CGI in their lungs and spleen.

* Increase in 24-h footpad swelling due to specific antigen compared to footpad injected with equal volume of saline. Footpad values >0.31 mm are considered positive; this figure was obtained by calculating two standard deviations from the mean footpad increment of animals developing positive DH responses. DH was produced by injecting mice intravenously with 5×10^5 SRBC 4 weeks after BCG injection. Five days later, animals were challenged in one rear footpad with 10^8 SRBC and in the other hind footpad with an equal volume of saline. Increases in footpad thickness were quantified using a pressure-sensitive caliper (Dyer).

† 33/124 (26.6%) is not different from an expected percentage of 25 ($P > 0.5$) predicted in the F_2 generation for a recessive trait controlled by a single gene.

Table 2 Strain survey of BCG-induced suppression of delayed hypersensitivity to sheep red blood cells

Mouse strain	BCG injected	*H–2*	*Igh*	Footpad increment*
C57BL/6	+	*b*	*b*	0.16 ± 0.02†
	–			0.60 ± 0.06
CBA	+	*k*	*a*	0.77 ± 0.06
	–			0.57 ± 0.06
C57BL/10	+	*b*	*b*	0.17 ± 0.06†
	–			0.68 ± 0.09
C57BR	+	*k*	*a*	1.12 ± 0.16
	–			0.41 ± 0.05
C57L	+	*b*	*a*	1.16 ± 0.09
	–			1.24 ± 0.26
BALB/c	+	*d*	*a*	0.51 ± 0.08
	–			0.46 ± 0.05
BALB/c.*Igh*^{b}	+	*d*	*b*	0.18 ± 0.02†
	–			0.60 ± 0.06
C3H.SW	+	*b*	*a*	0.48 ± 0.06
	–			0.41 ± 0.04
B10.BR	+	*k*	*b*	0.14 ± 0.03†
	–			0.59 ± 0.08

* See Table 1 legend.

† Negative footpad responses.

In summary, BCG-induced suppression of DH to SRBC is under genetic control. It is recessive, autosomal, due to a single (or closely linked group of) gene, and controlled by genes linked to the *Igh* complex.

Macrophage-mediated suppression is caused by a variety of agents, is usually not immunospecific, and suppresses a variety of immune responses[6–8]. Its overall role in immunoregulation is not understood. In our system, this type of suppression occurs in conjunction with intense CGI in the lungs and spleen. In another study, we have shown that adherent spleen cells from BCG-treated B6 mice produce soluble factors *in vitro* that suppress the production of CGI in the lungs and spleen[9]. Therefore, macrophage-mediated suppression may be important in the control of inflammation. Further studies are needed, however, to test this hypothesis.

The significance of linkage to *Igh* allotypes is not yet clear. In some studies, macrophage-mediated suppression is controlled by T lymphocytes[10,11], and several T cell-related immunological phenomena are linked to the *Igh* allotype[12,13], which may be related to V_H receptors on T cells. Studies are in progress in our laboratory to test these possibilities.

This research was supported by funds from HL22301, HL15389 and the Veterans Administration. We acknowledge the BALB/c.Igh^b mice from Dr Noel Warner, a gift of anti-*Igh* serum by Dr Mel Bosma, and the editorial assistance of Ms Mavis Baurle and Peggy LaSusa.

Received 19 February; accepted 27 October 1980.

1. Galindo, B. & Myrvik, Q. N. *J. Immun.* **105,** 227–237 (1970).
2. Allen, E. M., Moore, V. L. & Stevens, J. O. *J. Immun.* **119,** 343–347 (1977).
3. Allen, E. M. & Moore, V. L. *J. reticuloendothel. Soc.* **26,** 349–356 (1979).
4. Schrier, D. J., Allen, E. M. & Moore, V. L. *Fedn Proc.* **38,** 1174 (1979).
5. Moore, V. L., Schrier, D. J., Sternick, J. L. & Allen, E. M. in *Basic and Clinical Aspects of Granulomatous Inflammation* (eds Boros, D. L. & Yoshida, T.) (Elsevier, New York, in the press).
6. Coulis, P. A., Lewert, R. M. & Fitch, F. W. *J. Immun.* **120,** 1074–1076 (1978).
7. Klimpel, G. R. & Henney, C. S. *J. Immun.* **120,** 563–569 (1978).
8. Kirchner, H., Chused, T. M., Herberman, R. B., Holden, H. T. & Lavrin, D. H. *J. exp. Med.* **139,** 1473–1487 (1974).
9. Allen, E. M., Calvanico, N. J. & Moore, V. L. *Fedn Proc.* **38,** 1174 (1979).
10. Tadakuma, T. & Pierce, C. W. *J. Immun.* **120,** 481–486 (1978).
11. Ptak, W., Zembala, M. & Gershon, R. K. *J. exp. Med.* **148,** 424–434 (1978).
12. Eardley, D. D., Shen, F. W., Cantor, H. & Gershon, R. K. *J. exp. Med.* **150,** 44–50 (1979).
13. Weinberger, J. Z., Greene, M. I., Benacerraf, B. & Dorf, M. E. *J. exp. Med.* **149,** 1336–1348 (1979).

Chronic immune stimulation is required for Moloney leukaemia virus-induced lymphomas

John C. Lee & James N. Ihle

Biological Carcinogenesis Program, Frederick Cancer Research Center, Frederick, Maryland 21701

C-type viruses are known to be aetiologically related to naturally occurring leukaemia in a variety of species[1], although the mechanisms of transformation are largely unknown. The long latency periods, requirement for an acute viraemia and monoclonality of the tumours[2,3] distinguish leukaemogenesis from neoplasias induced by the acute transforming viruses and suggest an indirect mechanism. Consistent with this is the lack of *in vitro* transforming activity of replication-competent leukaemogenic viruses[4]. We have shown previously that the presence of T cells which proliferate *in vitro* in response to viral antigens is uniquely associated with the conditions leading to leukaemia. Based on these observations we have hypothesized that chronic immune stimulation is required for leukaemogenesis. We now demonstrate that CBA/N mice, when inoculated as newborns with Moloney leukaemia virus (MoLV), develop an acute viraemia but do not develop leukaemia or have detectable T-cell responses against the virus. This supports the hypothesis that chronic immune stimulation is essential for leukaemogenesis.

Previous studies have shown that MoLV can induce leukaemia in a variety of strains of mice[5]. However, as illustratedin Fig. 1, CBA/N mice failed to develop leukaemia after inoculation with a MoLV preparation which rapidly induced leukaemia in BALB/c mice. This result was surprising because the closely related CBA/J strain has been shown to be as susceptible to MoLV as BALB/c mice[5]. We therefore examined the virological and immunological status of MoLV-inoculated CBA/N mice.

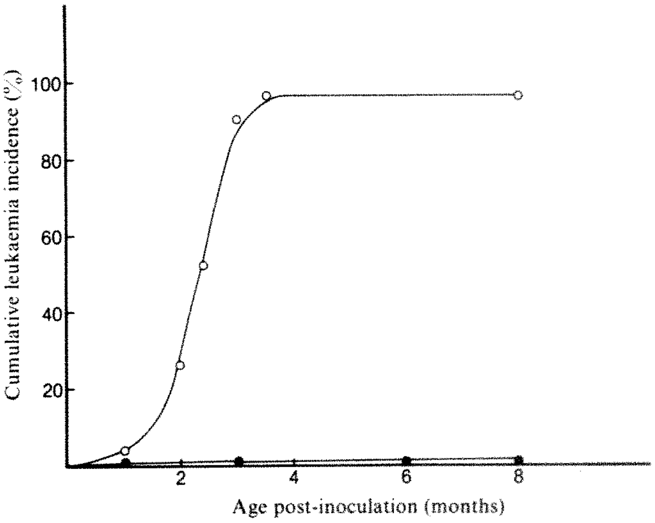


Fig. 1 Cumulative leukaemia incidence in MoLV-inoculated BALB/c or CBA/N mice. Groups of 20 BALB/c (○) or CBA/N (●) mice were inoculated as newborns with MoLV. The mice were reared in specific pathogen-free conditions. Moribund mice were killed and assessed histopathologically for the presence of lymphomas. The lymphomas in BALB/c mice were both of thymic and splenic origin, as previously described[23].

The ability of the virus to establish viraemia is known to be essential for leukaemogenesis[6,7]. To obtain a quantitative measure of virus replication we examined spleen extracts from inoculated mice for the viral proteins gp71, p30 and p12 using competition radioimmunoassays. As shown in Table 1, these proteins were readily detectable in CBA/N mice at 2 or 4 months post-inoculation. For comparison, the range of values obtained with MoLV-inoculated BALB/c or CBA/J mice are also shown. At 2 or 4 months post-inoculation, there were minimal differences in viral protein concentrations among the various strains. One important parameter which is known to be inversely related to leukaemia and viraemia is the presence of antibody against the virus[6]. Therefore, we examined the above mice for the presence of free antibody against MoLV. As shown in Table 1 none of the mice examined had detectable antcodies. These data demonstrate that the absence of leukaemia in CBA/N mice is not due to an inability of MoLV to replicate and establish viraemia.

It is well established that CBA/N mice differ from CBA/J mice in an immune response function and fail to develop immune responses to certain antigens early in life[8]. This deficiency has been shown to be due to a X-linked recessive locus[9]. Because our previous results demonstrated that the presence of T cells which respond *in vitro* by proliferation against viral proteins is correlated with leukaemogenesis[6,10], we compared this cellular immune response of MoLV-inoculated BALB/c, CBA/J and CBA/N mice. Splenic lymphocytes were obtained from individual mice, partially purified by passage through nylon wool columns as previously described[10] and assayed in a blastogenesis assay for proliferation in response to MoLV gp71. The responses of lymphocytes from individual BALB/c, CBA/J and CBA/N mice to various antigen concentrations are shown in Fig. 2. Lymphocytes from normal BALB/c, CBA/J and CBA/N did not respond to MoLV gp71. However, as previously demonstrated[10] and shown in Fig. 2, lymphocytes from preleukaemic BALB/c and CBA/J mice

Table 1 Levels of MoLV proteins and antibody against MoLV in inoculated BALB/c, CBA/J and CBA/N mice

Strain	Age (months)	Protein (ng per mg) gp71	p30	p12	Antibody titre
BALB/c	1	100–500	250–900	3–7	<1:20
BALB/c	2	300–1,400	400–1,100	10–50	<1:20
CBA/N	2	400–800	500–700	5–20	<1:20
BALB/c	3	900–1,700	700–1,500	60–70	<1:20
CBA/N	4	500–850	500–650	15–20	<1:20
CBA/J	4	1,500–1,900	1,000–1,200	30–50	<1:20

BALB/c, CBA/N and CBA/J mice were inoculated with MoLV at birth. At the indicated ages post-inoculation three to five individual mice were killed and used to determine the levels of antibody and viral proteins. The levels of viral proteins were determined using splenic extracts in competition radioimmunoassays for MoLV gp71, p30 and p12 as previously described[21]. The values represent the range of viral protein concentrations observed. The presence or absence of antibody in the serum was determined using ^{3}H-leucine-labelled intact MoLV virions in a radioimmune precipitation assay as previously described[22].

responded *in vitro* to MoLV gp71 with typical antigen concentration dependence. Previous studies[10] have shown this proliferative response to be dependent on a Thy-1.2$^+$, Lyt-1$^+$,2$^-$ T-cell subpopulation. In distinct contrast, however, lymphocytes from viraemic CBA/N mice had no detectable responses against MoLV gp71. Therefore, viraemic CBA/N mice seemed to be unique in lacking T cells specific for MoLV gp71.

We have also demonstrated that preleukaemic BALB/c mice have lymphocytes which respond *in vitro* to the internal viral protein p12 (ref. 10), although less consistently than to MoLV gp71. As shown in Fig. 2, lymphocytes from both preleukaemic BALB/c and CBA/J mice responded to p12 and showed typical antigen titration curves. Normal control BALB/c or CBA/J mice did not respond. The response to p12 in preleukaemic BALB/c and CBA/J mice was also found to depend on a Thy-1.2$^+$, Lyt-1$^+$,2$^-$ T cell (data not shown), whereas the MoLV-inoculated CBA/N mice had no detectable responses to p12.

We next examined individual MoLV-inoculated BALB/c, CBA/J and CBA/N mice at various ages. The results are summarized in Fig. 3, which presents the stimulation indices (SIs) against MoLV gp71 or p12. When lymphocytes from uninoculated BALB/c mice were examined, none had a significant blastogenic response (SI > 2.0) against gp71 or p12. In contrast, the majority of MoLV-inoculated BALB/c mice (86%) responded to gp71, whereas ~41% responded to p12. Interestingly, only one mouse (indicated by the solid circles in Fig. 3) failed to respond to either antigen. Using lymphocytes from individual MoLV-inoculated CBA/J mice, all the mice responded to MoLV gp71 *in vitro*, whereas five of the seven mice responded to MoLV p12. In distinct contrast, however, most of the CBA/N mice did not respond to either gp71 or p12. Only one mouse (indicated by the solid circles) had lymphocytes which responded at low levels to both MoLV gp71 and p12. These data demonstrate that MoLV-inoculated viraemic CBA/N mice are distinct from the other strains and generally lack a T-cell response against the viral proteins.

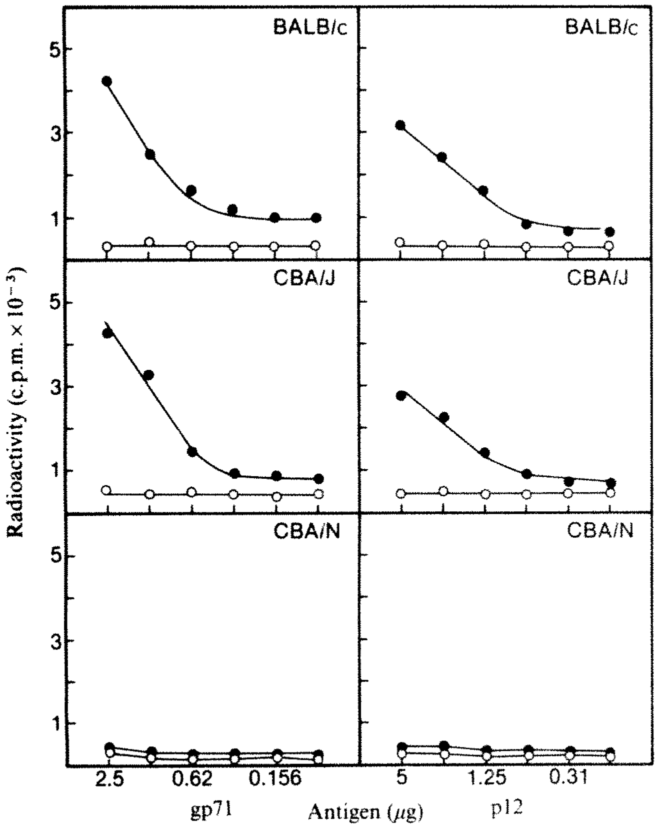


Fig. 2 T-cell proliferative responses of BALB/c, CBA/J and CBA/N mice to MoLV gp71 and p12. Lymphocytes were obtained from either 6-week-old normal mice (○) or mice inoculated as newborns with MoLV (●). An enriched T-cell preparation was obtained by nylon wool column fractionation and assayed in a microblastogenesis assay as previously described[24]. The lymphocytes were stimulated *in vitro* using purified MoLV gp71 (left panels) or p12 (right panels) at the indicated concentrations. The purification techniques have been previously described[24] and the antigens used were homogeneous as assessed by SDS-polyacrylamide gel electrophoresis. After 3 days of incubation the cells were pulse-labelled with ^{3}H-thymidine for 8 h and the incorporated radioactivity determined. Samples were done in triplicate and the individual samples varied less than 10% from the mean.

Our results describe two unique characteristics of the response of CBA/N mice to MoLV, that is, the lack of leukaemia and the lack of T cells which respond to MoLV proteins. The lack of leukaemia in viraemic mice is quite unusual. Gisselbrecht[5] has reported that DBA/2 mice fail to develop leukaemia by 6 months of age after MoLV inoculation at 4 weeks of age. However, we have observed that inoculated newborn DBA/2 mice develop leukaemia with a cumulative incidence of 85% by 8 months after inoculation (unpublished data). Interestingly, of the DBA/2 mice examined in those studies all hj T-cell proliferative responses to MoLV gp71. Therefore, we believe that CBA/N represents the only mouse strain in which MoLV-induced viraemia is not associated with leukaemia.

Although several genetic factors are known to influence MoLV-induced leukaemogenesis[5], most of these factors phenotypically decrease the level of viraemia. Our results clearly demonstrate that there were no significant differences in viraemia between the susceptible BALB/c and CBA/J strains and the resistant CBA/N mice. Another possible reason for the absence of leukaemia is the lack of an appropriate recombinant oncogenic virus. It has been suggested that spontaneous leukaemias in AKR mice require recombination to generate oncogenic MCF viruses[11,12]. In MoLV-inoculated CBA/N mice this seems unlikely for two reasons. First, cloned MoLV is oncogenic in rats and only the parental virus is found in most lymphomas (ref. 2 and D. Steffen, personal communication); therefore, recombination does not seem to be required. Second, the stock of MoLV used in these studies is an uncloned animal passaged stock which contains recombinant viruses similar to those previously described[13]. Nevertheless, the oncogenic potential and host range characteristics of viruses recovered from CBA/N mice are now being evaluated.

Our results demonstrate that viraemic CBA/N mice lack lymphocytes that respond to MoLV proteins whereas such lymphocytes are found in viraemic mice of other strains[6,10]. CBA/N mice are known to differ from CBA/J mice in their ability to respond to certain antigens[8]. However, the immune deficiency of CBA/N mice has only been shown to involve thymus-independent antigens, whereas our data suggest the absence of a T-cell function. One important distinction in our studies, however, is that MoLV was given to newborn mice, whereas other studies have examined the effect of immunizing adult mice. Therefore, the lack of responses may be related to the X-linked immune response gene and is manifested as tolerance. Preliminary genetic experiments support this in that the lack of a response to viral antigens is X linked.

The essential question is whether the absence of a detectable T-cell response against MoLV is responsible for the lack of leukaemia in CBA/N mice. Although the present results suggest this, appropriate genetic crosses will be required to confirm the relationship. If the lack of an immune response is involved, several possible explanations exist. McGrath and Weissman[14] presented evidence that T cells with appropriate virus-specific receptors were required for leukaemogenesis. If these receptors represent the immune specific receptors of T cells, our results demonstrate that CBA/N mice lack this subpopulation of T cells. Second, it is conceivable that an immune response is required to generate appropriate target-cell populations for virus infection and/or transformation. In a somewhat analogous manner we have previously hypothesized that chronic immune stimulation is required for leukaemogenesis. The basic

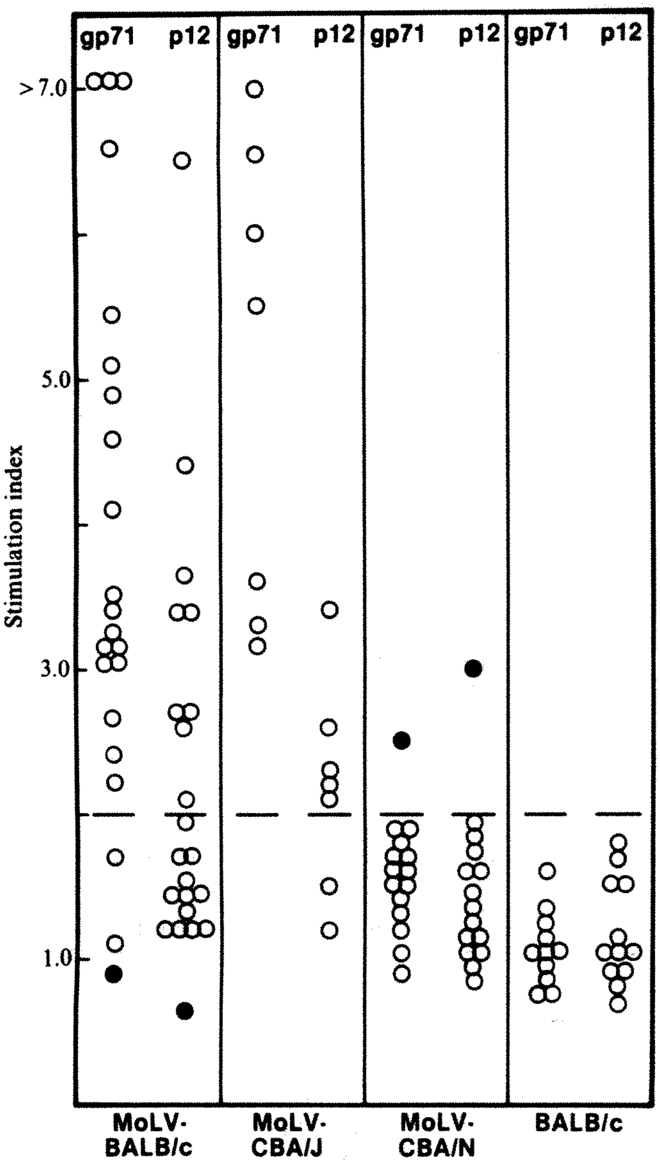


Fig. 3 T-cell proliferative responses of individual MoLV-inoculated or control BALB/c, CBA/N or CBA/J mice. Mice inoculated as newborns with MoLV or normal uninoculated BALB/c mice were used. Individual mice at 1–4 months of age were killed and splenic lymphocytes obtained and assayed in blastogenesis assays as described in Fig. 2 legend. Assays included the responses to MoLV gp71 (2.5 μg per well) or p12 (5 μg per well). All samples were tested in triplicate and the mean c.p.m. incorporated was determined. Individual samples generally varied less than 10% from the mean. The stimulation index presented is the ratio of c.p.m. of ^{3}H-thymidine incorporated in the presence of antigen to that in the absence of antigen. A stimulation index of 2.0 is considered a significant response. The solid circles indicate individual mice which either failed to respond to either antigen (BALB/c) or responded to both antigens (CBA/N), as discussed in the text.

components of this hypothesis are: (1) As a consequence of the establishment of an acute viraemia and expression of viral proteins, antigen reactive Thy 1.2$^+$, Lyt-1$^+$,2$^-$ T cells are continually recruited. (2) In response to antigen these T cells produce a variety of soluble lymphokines which influence the proliferation and differentiation of the immune system. Such factors include macrophage activating factors[15] which in turn cause the production of lymphocyte activating factor (LAF) or interleukin I (ref. 16). In addition, blastogenic factors are produced which promote proliferation of early T cells[17]. Other factors include those related to T-cell growth factors which promote thymocyte differentiation[18] and induction of differentiation of mature T cells[19]. (3) This chronic state of proliferation and differentiation increases the probability of spontaneous events leading to transformation. One likely 'error' commonly found associated with murine leukaemias that may be associated with transformation would be the generation of a trisomy in chromosome 15 (ref. 20). Irrespective of the validity of this interpretation, the results described here should provide important approaches to studying the mechanisms by which C-type viruses induce leukaemia.

We thank Jonathan Keller, Kathleen Bengali and Linda Rebar for technical assistance and Linda Fawley for help in preparation of the manuscript. This work was supported by NCI contract no. NO1-CO-75380, NIH.

Received 8 August; accepted 17 November 1980.

1. Tooze, J. *The Molecular Biology of Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1973).
2. Steffen, D. & Weinberg, R. A. *Cell* **15,** 1003 (1978).
3. Canaani, E. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1677 (1979).
4. Ihle, J. N. *et al.* in *Cold Spring Harb. Conf. Cell Proliferation* Vol. 7, 1049 (1980).
5. Gisselbrecht, S. *et al. Int. J. Cancer* **21,** 626 (1978).
6. Lee, J. C. & Ihle, J. N. *J. Immun.* **123,** 2351 (1979).
7. Lilly, F., Duran-Reynals, M. L. & Rowe, W. P. *J. exp. Med.* **141,** 882 (1975).
8. Amsbaugh, D. F., Hansen, C. T., Prescott, B., Barthold, D. R. & Baker, P. J. *J. exp. Med.* **136,** 931 (1972).
9. Berning, A. K., Eicher, E. M., Paul, W. E. & Scher, I. *J. Immun.* **124,** 1875 (1980).
10. Lee, J. C., Horak, I. & Ihle, J. N. *J. Immun.* (in the press).
11. Hartley, J. W., Wolford, N. K., Old, L. J. & Rowe, W. P. *Proc. natn. Acad. Sci. U.S.A.* **74,** 789 (1977).
12. Cloyd, M. W., Hartley, J. W. & Rowe, W. P. *J. exp. Med.* **151,** 542 (1980).
13. Fischinger, P. J., Nomura, S. & Bolognesi, D. P. *Proc. natn. Acad. Sci. U.S.A.* **72,** 5150 (1975).
14. McGrath, M. S. & Weissman, I. L. *Cell* **17,** 65 (1979).
15. Ng, A. K., Ames, R. S., McIntire, R. K. & Herberman, R. B. *Cancer Res.* **39,** 4887 (1979).
16. Farrar, J. J. & Koopman, W. J. in *Biology of the Lymphokines* (eds Cohen, S., Pick, E. & Oppenheim, J. J.) 325 (1979).
17. Enjuanes, L., Lee, J. & Ihle, J. N. *J. Immun.* (submitted).
18. Watson, J., Gillis, S., Marbrook, J., Mochizuki, D. & Smith, K. A. *J. exp. Med.* **150,** 849 (1979).
19. Ihle, J. N., Pepersack, L. & Rebar, L. *J. exp. Med.* (submitted).
20. Klein, G. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2442 (1979).
21. Fischinger, P. J. *et al. Virology* **90,** 241 (1978).
22. Ihle, J. N. & Joseph, D. R. *Virology* **87,** 287 (1978).
23. Pepersack, L., Lee, J. C., McEwan, R. & Ihle, J. N. *J. Immun.* **124,** 279 (1980).
24. Enjuanes, L., Lee, J. C. & Ihle, J. N. *J. Immun.* **122,** 665 (1979).

Phagocytosis requires repeated triggering of macrophage phagocytic receptors during particle ingestion

Denise R. Shaw & Frank M. Griffin Jr

Departments of Microbiology and Medicine, University of Alabama in Birmingham, Birmingham, Alabama 35294

The phagocytic functions of polymorphonuclear leukocytes and mononuclear phagocytes are crucial elements in host defence against a variety of invading microorganisms. Phagocytosis is a highly selective process, requiring specific interactions between the surface of the particle to be ingested and the plasma membrane of the phagocytic cell. The phagocyte can therefore discriminate between 'ingestible' and 'non-ingestible' particles even when the different particles are located in close proximity on the plasma membrane of the phagocytic cell[1]. Previous work has demonstrated that these specific interactions between particle and phagocyte are required not only for the initiation of phagocytosis—that is, attachment of the particle to the phagocytic cells and generation of a signal to trigger phagocyte pseudopod extension—but also for the subsequent progression of pseudopods over the entire surface of the particle[2,3]. We present evidence here that the continued interactions between phagocyte plasma membrane receptors and particle-bound ligands do not function merely to direct otherwise random phagocyte pseudopod movement, but instead are required for the repeated generation of intracellular phagocytic signals during the entire ingestion process.

All plasma membrane receptors for immunological ligands promote the binding of ligand-coated particles to the phagocytic cell surface; however, only certain receptors are able to generate a phagocytic signal, and this ability depends on the physiological state of the cell. Mouse peritoneal macrophages possess well characterized plasma membrane receptors of both phagocytic signal generating and non-generating classes. Receptors for the Fc portion of immunoglobulin G (IgG) molecules mediate both the attachment and the ingestion of particles coated with IgG (refs 4, 5). Receptors for the third component of complement (C3) mediate the attachment of particles coated with C3, but resident macrophages (R-MΦ) do not ingest the attached C3-coated particles[5-7]; however, inflammatory macrophages, such as those elicited by the intraperitoneal injection of Brewer's thioglycollate medium (TG-MΦ)[8], and resident macrophages treated *in vitro* with certain lymphokine preparations (LT-MΦ)[9] attach and ingest C3-coated particles.

Previous studies have indicated that the density of the immunological ligands on the surface of a particle can determine whether or not the particle will be phagocytosed. For example, particles coated with diminishing quantities of IgG are ingested by macrophages with progressively reduced efficiency; however, particles coated with inadequate amounts of IgG can be converted into particles which are readily ingested by R-MΦ after the addition of C3 ligands to the particle surface[10]. Other work has shown that if the phagocytosis-triggering ligands on an ingestible particle are sequestered on one hemisphere of the particle surface, that particle cannot be ingested[3]. From these previous experimental observations, we considered that there were two possible hypotheses for the role of phagocyte plasma membrane receptors in the process of particle ingestion: (1) Once a phagocytic cell has been triggered by a certain minimum number of phagocytic signals generated by appropriate ligand–plasma membrane receptor interactions, the phagocyte is committed to pseudopod extension of a degree sufficient to result in interiorization of the bound particle. Subsequent ligand–receptor engagement is required only to guide this pseudopod movement towards particle engulfment. Without such direction, fusion of the leading edges of pseudopods around the attached particle would be a random, inefficient process. (2) The engagement of phagocytic signal-generating plasma membrane receptors by the corresponding ligands triggers the phagocytic cell to produce only a limited amount of pseudopod extension. Unless such movement results in additional engagement of signal-generating receptors, particle engulfment will cease. Ligand–receptor interactions which do not generate phagocytic signals may function to facilitate each increment of pseudopod movement over the surface of the particle and thereby increase the efficiency of signal-generating receptor binding to additional ligands on the particle, but they cannot by themselves complete the process of particle engulfment.

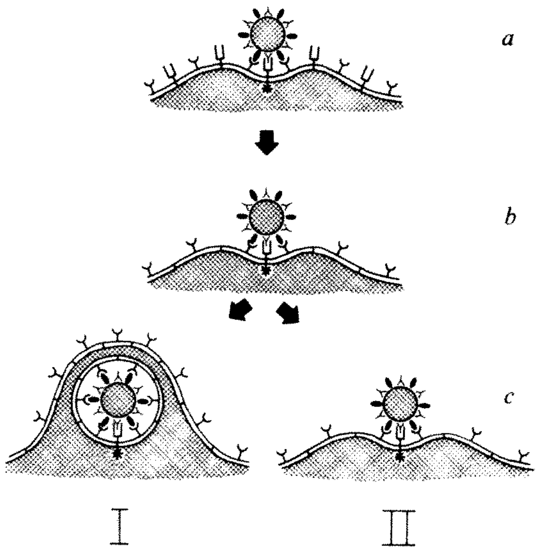


Fig. 1 Schematic representation of the experiment. *a*, IgG + C3-coated erythrocytes were bound to macrophage plasma membranes at 4 °C through the interactions of particle-bound IgG (Y) and C3 (●) ligands with the macrophages' Fc (Ψ) and C3 receptors (Y), respectively. *b*, Fc receptors located on the remainder of the macrophages' surfaces were functionally removed by treatment with anti-macrophage serum at 4 °C. *c*, Macrophages were then incubated at 37 °C for 30 min and examined for erythrocyte ingestion. If the initial phagocytic signal-generating (*) ligand–receptor (IgG–Fc) engagement triggers sufficient pseudopod extension, R-MΦ should complete erythrocyte ingestion through the 'passive' binding of C3 receptors to C3 ligands on the remainder of the erythrocyte surface (result I). However, if repeated phagocytic signal-generating ligand–receptor interactions are required for each increment of pseudopod extension, then IgG + C3-coated erythrocytes should not be ingested by R-MΦ (result II).

To discern which of the two hypotheses best describes the mechanism of phagocytosis, we have performed the experiment schematically shown in Fig. 1. First, macrophages were allowed to attach, through their Fc and C3 receptors, particles coated with both IgG and C3, in conditions which prevented particle

Table 1 Attachment and ingestion of erythrocytes by macrophages

Erythrocytes coated with:	Attachment index ± s.e.m.			Phagocytic index ± s.e.m.		
	R–MΦ	TG-MΦ	LT-MΦ	R-MΦ	TG-MΦ	LT-MΦ
			Control MΦ			
1. Nothing	8.1 ± 1.6 (3)	20 ± 9.1 (3)	40 ± 17 (3)	0 (3)	7.4 ± 4.7 (5)	7.9 ± 6.4 (3)
2. C3	1,200 ± 100 (3)	940 ± 28 (2)	840 ± 140 (3)	9.8 ± 6.8 (3)	490 ± 180 (2)	350 ± 85 (3)
3. IgG	510 ± 139 (4)	1,100 ± 290 (4)	1,100 ± 200 (3)	300 ± 59 (5)	700 ± 230 (4)	670 ± 96 (4)
4. IgG + C3	710 ± 280 (4)	970 ± 160 (4)	2,000 ± 310 (3)	440 ± 110 (5)	610 ± 140 (4)	1,000 ± 54 (4)
			MΦ treated with anti-macrophage serum			
5. IgG	520 ± 100 (7)	960 ± 250 (4)	1,200 ± 280 (4)	13 ± 4.2 (8)	29 ± 9.5 (4)	53 ± 6.5 (8)
6. IgG + C3	750 ± 130 (7)	900 ± 320 (5)	1,700 ± 170 (4)	19 ± 9.5 (8)	460 ± 48 (5)	310 ± 24 (8)

Macrophages from non-treated mice (R-MΦ) or from thioglycollate-injected mice (TG-MΦ) were cultured for 48 h before the assay. Some macrophages from non-treated mice were cultured in medium containing the previously described lymphokine (LT-MΦ) which enhances macrophage C3 receptor function[9]. After attachment of erythrocytes to macrophages, some cultures were treated with anti-macrophage serum at a concentration that reduced Fc receptor-mediated phagocytosis by 90% but was not toxic to macrophages. Results are expressed as attachment or phagocytic indices, calculated by multiplying the percentage of macrophages attaching or ingesting one or more erythrocytes by the mean number of erythrocytes attached or ingested per macrophage. Numbers in parentheses indicate the number of determinations, each determination being made on 100 macrophages. The macrophage cultures were relatively homogeneous, and attachment and phagocytic indices are representative of the population as a whole. Thus, small indices (<60) were obtained only when less than 20% of macrophages attached or ingested a few erythrocytes each, whereas large indices (>300) were never the result of a minority of macrophages attaching or ingesting large numbers of erythrocytes but were obtained only when greater than 65% of the population participated in the function assessed.

ingestion (Fig. 1*a*). Non-engaged macrophage Fc receptors were specifically blocked without affecting the preformed receptor interactions (Fig. 1*b*). Cultures were then restored to conditions permissive for ingestion (Fig. 1*c*). The protocol used was intended to provide an initial phagocytosis trigger in the form of the engagement of macrophage Fc receptors by particle-bound IgG. However, completion of particle ingestion, if it occurred, would have to proceed through macrophage C3 receptor binding to particle-bound C3, interactions which alone are unable to promote ingestion by R-MΦ.

Procedures for obtaining monolayer cultures of R-, TG- and LT-MΦ and for preparing sheep erythrocytes (particles) coated with IgG, C3 or IgG + C3 have been described in detail elsewhere[1-3,7-9]. Erythrocytes were incubated with macrophage cultures at 4 °C for 30 min, conditions which allow the attachment but not the ingestion of ligand-coated particles[11]. Cultures were then treated for 60 min at 4 °C with the IgG fraction of a rabbit anti-mouse macrophage serum which has previously been shown to block specifically macrophage Fc receptor function without affecting macrophage C3 receptors[8,12]. Cultures were then transferred to 37 °C and an atmosphere of 95% air–5% CO_2, conditions which are permissive both for phagocytic signal transmission by IgG–Fc receptor bonds previously formed at 4 °C (ref. 3) and for particle ingestion[2,3]. After incubation for 30 min, cultures were prepared for phase contrast microscopy and scored for erythrocyte phagocytosis as previously described[1].

Control experiments (Table 1, lines 1–5) confirmed that non-coated erythrocytes were neither bound nor ingested by any macrophage preparation (line 1), that C3-coated erythrocytes were attached but not ingested by R-MΦ and were ingested by TG- and LT-MΦ (line 2), that erythrocytes coated with IgG (line 3) and those coated with IgG + C3 (line 4) were attached and ingested by all three types of macrophages, and that treatment with anti-macrophage serum did not affect the previously established attachment of IgG-coated erythrocytes, but did efficiently block the subsequent ingestion of these erythrocytes by all macrophage preparations (compare lines 3 and 5).

As expected, TG- and LT-MΦ were able to ingest IgG + C3-coated erythrocytes even after treatment with anti-macrophage serum (line 6, columns 5, 6). The phagocytic indices were nearly identical to those obtained when untreated TG- and LT-MΦ were incubated with erythrocytes coated with C3 alone (line 2), in agreement with previous observations that anti-macrophage serum treatment completely blocks macrophage Fc receptors but does not impair macrophage C3 receptor function[8,12]. In contrast, R-MΦ treated with anti-macrophage serum were unable to ingest IgG + C3-coated erythrocytes (line 6, column 4).

The inability of R-MΦ to ingest IgG + C3-coated erythrocytes supports hypothesis (2). These erythrocytes presented sufficient C3 ligands so that, even after macrophage Fc receptors had been blocked, they could be ingested by TG- and LT-MΦ, macrophages for which C3 receptor engagement is capable of generating phagocytic signals. However, in similar conditions, R-MΦ were unable to ingest these IgG + C3-coated particles, even though the initial interaction of particle with macrophage could occur by phagocytic signal-generating ligand–receptor (IgG–Fc) engagements and even though the R-MΦ could efficiently attach particles through their C3 receptors.

We thank Ms Peggy M. Rains for technical assistance and Dr Roy Curtiss III for very helpful criticism of the manuscript. This work was supported by American Cancer Society grant IM-173. D.R.S. is supported by NIH Predoctoral Training grant GM-07164. F.M.G. is the recipient of NIH Research Career Development Award AI-00135.

Received 22 August; accepted 13 November 1980.

1. Griffin, F. M. & Silverstein, S. C. *J. exp. Med.* **139,** 323–336 (1974).
2. Griffin, F. M., Griffin, J. A., Leider, J. E. & Silverstein, S. C. *J. exp. Med.* **142,** 1263–1282 (1975).
3. Griffin, F. M., Griffin, J. A. & Silverstein, S. C. *J. exp. Med.* **144,** 788–809 (1976).
4. Berken, A. & Benacerraf, B. *J. exp. Med.* **123,** 119–144 (1966).
5. Mantovani, B., Rabinovitch, M. & Nussenzweig, V. *J. exp. Med.* **135,** 780–792 (1972).
6. Lay, W. H. & Nussenzweig, V. *J. exp. Med.* **128,** 991–1009 (1968).
7. Griffin, F. M., Bianco, C. & Silverstein, S. C. *J. exp. Med.* **141,** 1269–1277 (1975).
8. Bianco, C., Griffin, F. M. & Silverstein, S. C. *J. exp. Med.* **141,** 1278–1290 (1975).
9. Griffin, J. A. & Griffin, F. M. *J. exp. Med.* **150,** 653–675 (1979).
10. Ehlenberger, A. G. & Nussenzweig, V. *J. exp. Med.* **145,** 357–371 (1977).
11. Rabinovitch, M. *Expl Cell Res.* **46,** 19–28 (1967).
12. Holland, P., Holland, N. H. & Cohn, Z. A. *J. exp. Med.* **135,** 458–475 (1972).

Nonjunctional acetylcholine receptor channel open time decreases during development of *Xenopus* muscle

R.W. Kullberg

Biology Department, University of Alaska, Anchorage, Anchorage, Alaska 99504

P. Brehm & J. H. Steinbach

The Salk Institute, PO Box 85800, San Diego, California 92138

The mean channel open time of junctional acetylcholine receptors (AChRs) decreases during development of mammalian skeletal neuromuscular junctions[1-3]. The junctional AChRs of newborn rats have long mean channel open times of about 4.5 ms and in this respect are similar to nonjunctional AChRs of adult denervated muscle[3-5]. Mature neuromuscular junctions have channels with shorter mean open times of about 1 ms, whereas junctions at intermediate stages seem to have a mixture of channels with short and long open times[1-3]. These observations, taken together with findings on receptor distribution during development[6,14], suggest that the establishment of nerve–muscle contact induces an aggregation of embryonic AChRs with long mean channel open times at the point of contact and that these embryonic junctional receptors are then replaced, modified or subjected to local membrane changes leading to a discrete reduction in channel open time. However, the fate of those embryonic nonjunctional receptors which are not aggregated at the newly formed neuromuscular junction is unknown. Are they an unchanging population, or do the changes in kinetic behaviour which occur at the junctional regions also take place at other regions of the muscle membrane? The kinetics of nonjunctional receptors in mammalian skeletal muscle have not been studied during development. We have now, however, examined this question in amphibian myotomal muscle, where nonjunctional receptors are present in substantial numbers throughout development, making it possible to measure the mean open time of both classes of AChR channel. We observed a shortening of mean channel open time of junctional AChRs, from about 3 ms to 1 ms, and an identical change in mean channel open time of AChRs at nonjunctional regions of membrane.

The experiments were performed on myotomal muscles of *Xenopus laevis* embryos and tadpoles. Procedures of dissection and staging have been described previously[7]. The muscle cells extend the length of a myotome and insert at myocommata where the neuromuscular junctions are located. Focal synaptic potentials may be recorded with an extracellular pipette positioned at the myocommata[8], and staining with fluorescent-labelled α-bungarotoxin reveals extensive receptor accumulations at the ends of the muscles[9]. Muscle cells ranged in length from 100 μm in the earliest stages studied to 200 μm in the latest stages.

Recordings of ACh-induced extracellular noise were obtained from muscles using the method developed by Schuetze, Frank and Fischbach[10]. An extracellular pipette (3–7 μm o.d.) filled with bathing solution (see Fig. 1 legend) containing either 25 or 50 μM ACh was pressed against the muscle to record agonist-induced noise. Recordings were

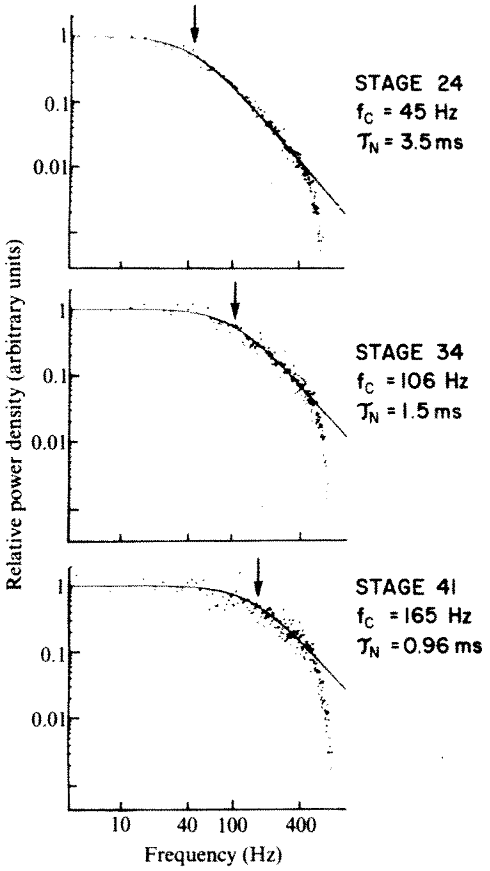


Fig. 1 Sample spectra for fluctuations recorded from nonjunctional regions of *Xenopus* myotomal muscle cells at three different developmental stages. The spectra are plotted logarithmically and the power is expressed in arbitrary units. Data were bandpass filtered at 0.5–500 Hz and digitized at 0.49-ms intervals. Spectra were computed with a fast Fourier-transform routine using a Digital PDP 11/34 computer. The spectra shown have had background spectra subtracted (this correction resulted in a change of less than 10% in the power at frequencies less than approximately twice the half-power frequency). The background spectra were obtained from records in which the pipette was not pressed against a cell. Single lorentzian curves are shown, fitted to the data by eye. A lorentzian curve is described by the equation $S(f) = S_0/[1+(f/f_c)^2]$, where S_0 is the zero frequency power, f_c is the half-power frequency and f is frequency. The recording solution contained 111 mM NaCl, 3 mM KCl, 1.8 mM $CaCl_2$ and 8 mM HEPES-NaOH buffer (*p*H 7.4). Records were made at room temperature (20–23 °C).

classified as 'nonjunctional' when the electrode tip was placed at the middle of the myotome, where no synapses are located, and 'junctional' when positioned at the myocommata. In the latter case, the occurrence of infrequent, miniature endplate currents (m.e.p.c.s) in some recordings confirmed the close proximity of the electrode tip to junctional regions. To avoid possible differences in myotomal development along the length of the tadpole, only the anterior myotomes were included in this study. Sealing of the ACh-containing pipette against the muscle membrane (sealing resistance 0.5–5 MΩ) was accompanied by an increase in voltage fluctuations and a mean negative voltage shift. This signal resulted from the activity of ACh-activated channels, as it was abolished by (+) tubocurarine (3×10^{-5}M) and was not observed when the electrode was either lacking ACh or pressed against epidermis.

Estimates of mean channel open time were obtained from spectral analysis of ACh-induced noise. Power density spectra were computed and fitted by single lorentzian curves (Fig. 1) which describe the theoretical spectra produced by randomly opening ACh-activated channels[11]. In conditions of low ACh concentration, the half-power frequency (f_c) of a spectrum of fluctuations recorded from receptors on adult amphibian muscle is determined only by the channel-closing rate constant and provides an estimate (τ_n) of the mean channel open time according to the equation $\tau_n = 1/(2\pi f_c)$ (ref. 11). At high concentrations of ACh, τ_n will underestimate the true mean channel open time, because the half-power frequency is then determined by both channel opening and closing rates[11,12]. To minimize the error in our estimate of mean channel open time, we have used the lowest possible ACh concentration which gave consistently satisfactory recordings. Because of dilution, the ACh concentration at the tip of the recording pipette was less than 25–50 μM. In most cases, a pipette was used until dilution rendered the ACh concentration so low that no further response could be recorded.

Spectral analysis of both junctional and nonjunctional AChR channel fluctuations was performed at developmental stages between 2 and 58 h after initial nerve contact. The power in the fluctuations extended to higher frequencies at later stages, yielding estimates of mean channel open times of the AChRs that are shorter in later developmental stages (Fig. 1). Values of τ_n were determined from spectra recorded at 101 sites (39 embryos). The values of τ_n measured at junctional regions decreased from 3.0±0.7 ms (mean ±s.d., $n = 8$) at stage 24 to 1.1±0.3 ms ($n = 7$) at stages 38–42. These values were not significantly different from the corresponding values for the nonjunctional AChRs, 2.7±0.9 ms ($n = 8$) at stage 24 and 1.2±0.3 ms ($n = 9$) at stages 38–42 (Fig. 2).

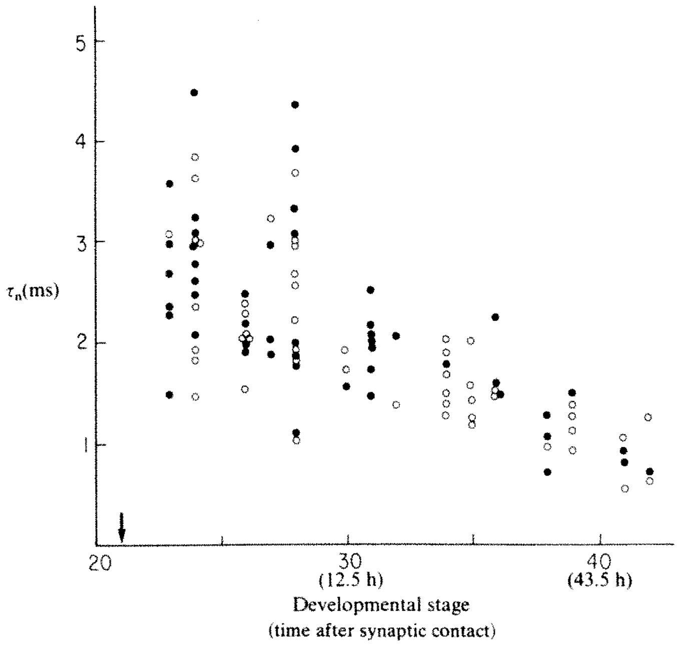


Fig. 2 The calculated τ_n for junctional (filled circles) and nonjunctional (open circles) AChRs are plotted as a function of developmental stage. The abscissa is also marked in terms of hours of development after the first signs of synaptic transmission[7,15]. The arrow indicates the time when the first synaptic activity may be recorded. Each point gives τ_n for a single recording site.

Developmental changes in junctional AChR channel kinetics were also examined by analysis of focally recorded synaptic currents (focal m.e.p.c.s). In conditions where the transient rise and fall of cleft concentration of transmitter is rapid relative to channel closing kinetics, the decay time constant of synaptic currents provides an independent estimate (τ_m) of the mean

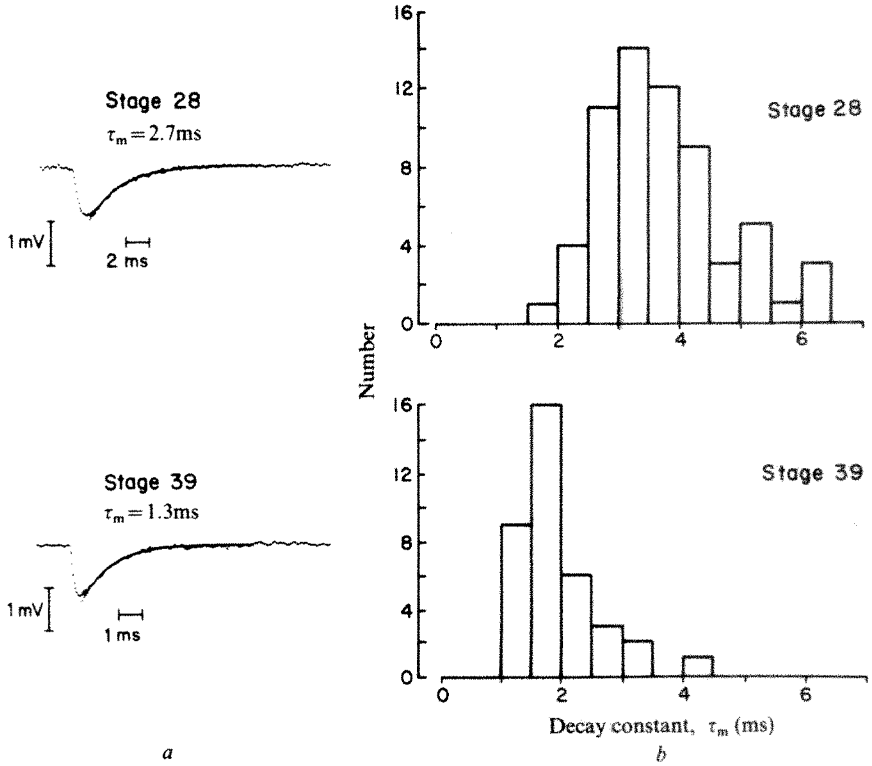


Fig. 3 *a*, Focal m.e.p.c.s recorded at stages 28 and 39. Recordings were made by positioning the tip of a Ringer-filled extracellular pipette at the myocommata. Focal m.e.p.c.s were bandpass filtered (0.5–2 kHz) and sampled on-line by computer. The individual m.e.p.c. decays were then fitted by the best least-squares exponential curve. These curves are shown superimposed on the m.e.p.c. decay phases. *b*, Histograms of decay constants of focal m.e.p.c.s at stages 28 and 39. Each histogram is based on a sample of focal m.e.p.c.s recorded at a single site.

channel open time[11,13]. In developing rat neuromuscular junctions, there is a fairly good correspondence between the time constant of m.e.p.c. decay and spectral estimates of mean channel open time[1–3]. We recorded samples of focal m.e.p.c.s at selected stages of development, and the decay phases were fitted by single exponential curves (Fig. 3*a*). At any single recording site, there was a wide scatter of estimated decay constants, many of which were longer than the spectral estimates of mean channel open time at the same stage (Fig. 3*b*). However, the shorter decay constants, as illustrated by the m.e.p.c.s in Fig. 3*a*, were similar to spectral estimates of the mean channel open time at each stage. We never observed focal m.e.p.c.s which had briefer time constants than the fastest spectral estimates of mean channel open time at the same stage. These facts suggest that in the developing myotomal neuromuscular junction there is a fraction of ACh release sites where the decay of synaptic current is limited by channel kinetics and not by the rate of disappearance of transmitter.

We were unable to obtain measurements of τ_n in animals beyond stage 42 using the focal noise technique, possibly because of a decrease in nonjunctional receptor density and an increase in connective tissue at the myocommata, which made the receptors less accessible to the ACh-containing electrode. We were, however, able to estimate τ_m from focal m.e.p.c.s. In stage 50 muscle, τ_m was 0.7 ± 0.1 (mean ± s.d., $n = 181$ focal m.e.p.c.s; 6 recording sites); this indicates that receptor kinetics continue to accelerate beyond our stage 42 noise measurements.

The acceleration of junctional AChR channel kinetics observed at the mammalian neuromuscular junction results from a discrete shift to shorter mean channel open times[1–3]. If this mechanism explained the shift in τ_n seen for *Xenopus* AChRs, the spectra at intermediate stages should be fitted by the sum of two lorentzian functions corresponding to the two populations of receptors. The values for τ_n would be expected to correspond, as in the case of rat muscle, to the slowest and fastest estimated values of mean channel open time at early and late stages. However, of 69 spectra recorded from intermediate stages (26–39), 50 spectra were satisfactorily fitted by eye with single lorentzian curves, as in Fig. 1, stage 34. The remaining spectra seemed to be fitted better with double lorentzian curves composed of slow and fast components. Further examination of high resolution spectra will be necessary to determine whether the changes in channel kinetics occur gradually or discretely.

The importance of the present study is the finding that a modification of AChR channel kinetics occurs in both junctional and nonjunctional AChRs in developing *Xenopus* myotomal muscle. The process of change in AChR channel kinetics does not seem to depend on the proximity of receptors to the site of nerve contact with muscle, but affects receptors over the entire surface membrane. The development of the mammalian skeletal neuromuscular junction is similar to that of the myotomal synapse in that nonjunctional receptors are present in significant numbers (150 μm^{-2})[14] during the period of transition from slow to fast kinetics of junctional receptors. Given the present observations of *Xenopus* embryonic receptors, it would be interesting to see if similar changes in the kinetics of mammalian embryonic nonjunctional receptors occur during neuromuscular junction development.

This work was supported by NIH and NSF grants to J.H.S. and a Muscular Dystrophy Association grant to R.W.K. P.B. was supported by a Muscular Dystrophy Association Postdoctoral Fellowship. We thank Drs Y. Kidokoro, S. Halegoua, V. Dionne and R. Jaffe for critical reading of this paper.

Received 11 August; accepted 15 October 1980.

1. Sakmann, B. & Brener, H. R. *Nature* **276**, 401–402 (1978).
2. Schuetze, S. M. & Fischbach, G. D. *Soc. Neurosci. Abstr., 8th a. Meet.* 374 (1978).
3. Fischbach, G. D. & Schuetze, S. M. *J. Physiol., Lond.* **303**, 125–137 (1980).
4. Dreyer, F., Muller, K.-D., Peper, K. & Sterz, R. *Pflügers Arch. ges. Physiol.* **367**, 115–122 (1976).
5. Sakmann, B. *Fedn Proc.* **37**, 2654–2659 (1978).
6. Anderson, M. J. & Cohen, M. W. *J. Physiol., Lond.* **268**, 757–773 (1977).
7. Kullberg, R. W., Cohen, M. W. & Lentz, T. L. *Devl Biol.* **60**, 101–129 (1977).
8. Kullberg, R. W., Mikelberg, F. S. & Cohen, M. W. *Devl Biol.* **75**, 255–267 (1980).
9. Anderson, M. J. & Cohen, M. W. *J. Physiol., Lond.* **237**, 385–400 (1974).
10. Schuetze, S. M., Frank, E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 520–523 (1978).
11. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–691 (1973).
12. Sakmann, B. & Adams, P. R. *Proc. 7th int. Congr. Pharmac.*, Paris, 1978 (ed. Jacob, J.) 81–90 (Pergamon, Oxford, 1979).
13. Colquhoun, D., Large, W. A. & Rang, H. P. *J. Physiol., Lond.* **266**, 361–395 (1977).
14. Bevan, S. & Steinbach, J. H. *J. Physiol., Lond.* **267**, 195–213 (1977).
15. Blackshaw, S. & Warner, A. *Nature* **262**, 217–218 (1976).

Interferon action—sequence specificity of the ppp(A2′p)$_n$A-dependent ribonuclease

Daniel H. Wreschner*, John W. McCauley*, John J. Skehel & Ian M. Kerr*

National Institute for Medical Research, Mill Hill,
London NW7 1AA, UK

The oligonucleotides pppA2′p5′A2′p5′A and related oligomers (2-5A) are synthesized by an enzyme that is widely distributed in a variety of cells, the activity of which varies with interferon treatment, growth and hormone status[1,2]. Because significant amounts of 2-5A have recently been detected in interferon-treated cells[3], it has been suggested that the oligonucleotides may be involved in interferon action and in the control of cell metabolism. In both intact cells[4,5] and cell-free systems[6–11] 2-5A has been shown to activate a ribonuclease. We report here investigations of the sequence specificities of the 2-5A-dependent ribonucleases in extracts of rabbit reticulocytes, mouse ascites tumour cells and human lymphoblastoid cells in conditions of partial digestion using terminally labelled RNA substrates. The enzymes cleaved on the 3′-side of UN sequences to yield UpNp terminated products. Cleavage was observed predominantly at UA and UU sequences.

Partial nuclease digestions in the presence and absence of tetramer 2-5A were done using rabbit reticulocyte lysates[12] with the 3′-terminally labelled small RNA of the La Crosse virus genome[13] and influenza virus genome RNAs 7 and 8 as substrates[14]. The products were compared by electrophoresis in polyacrylamide gels, with those obtained by specific limited chemical degradations of the RNAs[15]. For La Crosse virus sRNA, in the presence of 2-5A, specific cleavages were observed at UN sequences where UA and UU were preferred (Fig. 1*a*). In this experiment cleavage was obtained at nucleotides 32, 30, 29 and 34, the first of which is at a UA and the others at UU sequences. The precise nature of the products of digestion is not defined by these results and will be described below. Analysis of larger oligonucleotide products of digestion indicated cleavage at residues 54, 58, 69, 100 and 149, in all cases at UA sequences (Fig. 1*b*, Table 1*a*). Additional cleavages at nucleotides 59, 66, 73, 123 and 124 were at UU sequences and in the case of nucleotide 72, at a UG. These data are summarized in Table 1. Denaturation of the RNA by boiling before addition to the lysate did not alter the pattern of cleavage (Fig. 1*b*). In addition, identical results were obtained when 2-5A tetramer, pentamer or oligomers obtained from interferon-treated, virus-infected cells[3] were used for activation (data not shown).

The same specificity was observed using influenza virus RNAs 7 and 8 as substrates (Fig. 1*c*, *d*; Table 1). With RNA 7 the major sites of degradation were at nucleotides 17, 36 and 64 (Fig. 1*c*), all at UA sequences. Less extensive cleavage also occurred at nucleotides 23 and 49 (UU sequences), and 53 and 57 (UA sequences). For RNA 8 distinct cleavages were obtained at residues 19, 20, 21, 25 and 26. The major site of cleavage at

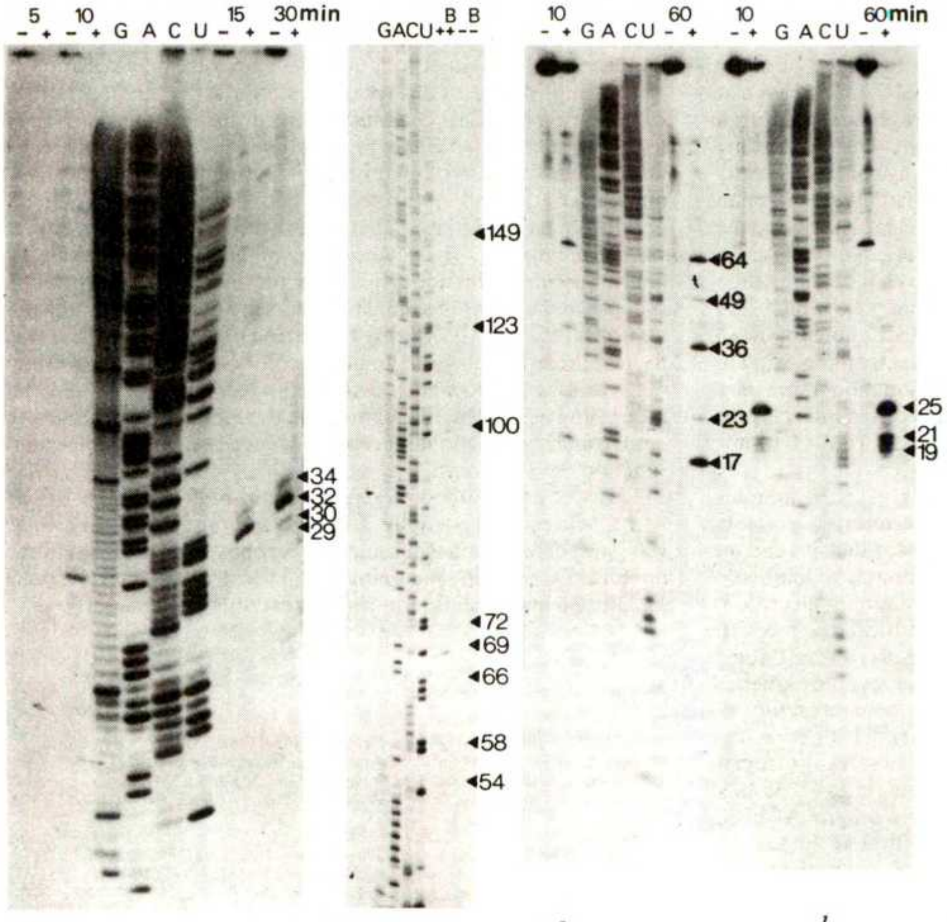


Fig. 1 Comparative electrophoretic analysis of the products of 2-5A-activated RNase activity and those obtained by limited chemical degradation. Virus RNAs were labelled at their 3′-termini with cytidine 3′5′-[5-^{32}P]biphosphate and T$_4$ RNA ligase[21] and purified by polyacrylamide gel electrophoresis[13,16]. Assays were performed using rabbit reticulocyte lysates in protein synthetic conditions[12] but in the absence of creatine phosphokinase. Nuclease digestions in the presence (+) or absence (−) of tetramer 2-5A (100 nM) were at 30 °C for up to 60 min. Reactions were terminated by cooling to 0 °C, glycerol was added to 20% w/v and aliquots (10 µl) were applied to polyacrylamide gels containing 7 M urea, 100 mM Tris-borate *p*H 8.3, and 2.5 mM EDTA. 3′-Labelled RNAs specifically degraded as described by Peattie[15] were applied to adjacent lanes, labelled GACU. *a*, Autoradiogram of a 20% polyacrylamide gel (40×40×0.05 cm) containing the products of cleavage of La Crosse sRNA after incubation for 5, 10, 15 and 30 min. *b*, Autoradiogram of a 12% polyacrylamide gel (100×20×0.05 cm) containing the digestion products of La Crosse sRNA after 10 min incubation. In the samples labelled B the RNA substrate was incubated in a boiling water bath for 1 min immediately before addition to the digestion mixtures. *c*, *d*, Autoradiograms of 12% polyacrylamide gels (40×40×0.05 cm) containing the digestion products of X31 influenza virus RNAs 7 (*c*) and 8 (*d*) after incubation for 10 and 60 min. The numbers indicate the positions of cleavage of the RNA molecules by the 2-5A-dependent nuclease.

---

* Present addresses: Department of Microbiology, Tel Aviv University, Ramat Aviv, Israel (D.H.W.); Department of Pathology, University of Cambridge, UK (J.W.M.); ICRF Laboratories, Lincoln's Inn Fields, London WC2A 3PX, UK (I.M.K.).

0028–0836/81/050414—04$01.00

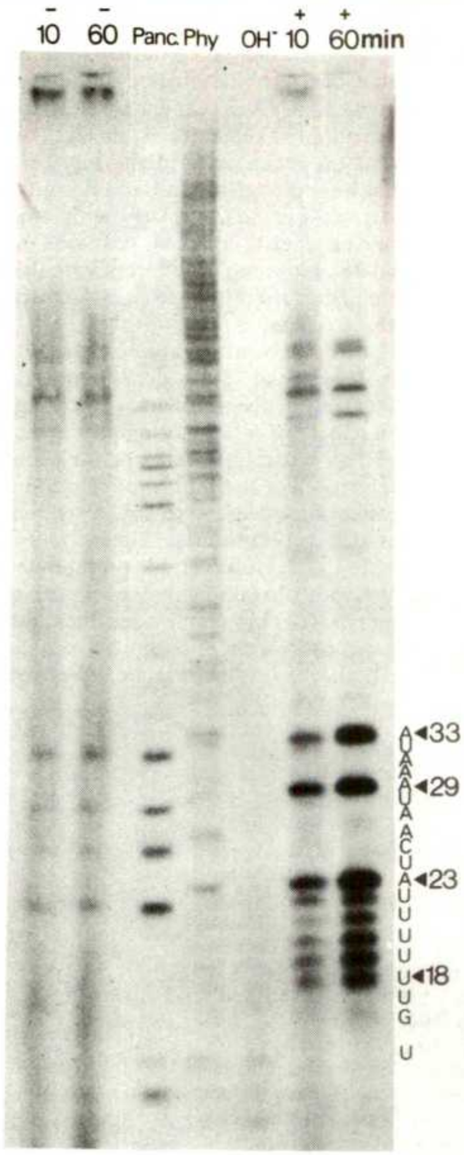


Fig. 2 Analysis of the products of digestion of 5′-labelled influenza virus RNA 8 by the 2-5A-activated RNase from rabbit reticulocytes. X31 influenza virus RNAs were labelled at their 5′-termini using polynucleotide kinase and [γ-^{32}P]ATP and separated as described elsewhere[16]. RNA 8 was incubated in the reticulocyte lysate in the presence (+) and absence (−) of 2-5A for 10 and 60 min as described in Fig. 1 legend. The products of digestion were separated on a 12% polyacrylamide gel (40 × 40 × 0.05 cm) in parallel with the products of partial alkaline hydrolysis (OH⁻) and partial digestions with pancreatic ribonuclease (Panc) and *Physarum polycephalum* nuclease M (Phy) as described elsewhere[13,22,23]. The relevant nucleotide sequence of RNA 8 is indicated.

position 25 was at a UA sequence, the other sites at UU sequences (Figs 1*d*, 3*b*; Table 1*c*).

With 3′-terminally labelled 18S ribosomal RNA from rabbit reticulocytes, cleavage was exclusively observed at UA and UU sequences (data not shown).

To define further the sites of cleavage by the 2-5A-activated RNase, the products obtained using 5′-terminally labelled influenza virus RNA 8 were examined. As shown in Fig. 2 and detailed in Table 1*d*, the products of partial digestion in the presence of 2-5A have identical electrophoretic mobilities to the oligonucleotides terminating at residues 18, 19, 20–23, 29 and 33 which were obtained after partial digestion with *Physarum* or pancreatic nuclease, or partial alkaline hydrolysis, all of which terminate in 3′ phosphates[16]. This indicates that all the products obtained from the reticulocyte lysate were phosphorylated at their 3′ termini and also that their 3′-terminal sequences were UpNp. These conclusions are consistent with the data presented in Fig. 1, where the electrophoretic mobilities of the 2-5A-activated nuclease digestion products were less than those of the oligonucleotides derived from chemical degradation, which contain 5′-terminal phosphates and from which the modified bases have been eliminated.

2-5A activates a ribonuclease and inhibits protein synthesis in a variety of cell-free systems[6–11]. To determine whether or not these enzymes have a common specificity, cell extracts were

Table 1 Nucleotide sequence* of the 3′- and 5′-labelled RNA substrates indicating sites of cleavage observed in the presence of 2-5A

3′-labelled RNAs

a La Crosse sRNA
10 20 30
3′ U C A U C A C A U G A G G U G A A C U U A U G A A A C U $\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$
U $\overset{\downarrow}{A}$ U $\overset{\downarrow}{U}$ U A A C A A C A A C U G A C A A A A A $\overset{\downarrow}{A}$ U G G $\overset{\downarrow}{A}$ $\overset{\downarrow}{U}$ U
C C C C U $\overset{\downarrow}{U}$ U A $\overset{\downarrow}{A}$ U A $\overset{\downarrow}{G}$ $\overset{\downarrow}{U}$ U C U C A C A C U A C A G C C U A
A A C C A C A A A A U A C U A C A G C G U A G U U G U C C A
C G $\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$ U A C C U A A A C U A G G A C U A C G U C C A U $\overset{\downarrow}{A}$ U
A C C U G A A G A C A C A A U U U U U A C G U C U U A G U G
A G U U G G A A C G A C G U C A A U C U A G A A G A A G G A
G U A C G

b X31 influenza RNA 7†
10 20 30
3′ U C G C U U U C G U C C A U C U $\overset{\downarrow}{A}$ U A A C U $\overset{\downarrow}{U}$ U C U A C U C
A G A A G $\overset{\downarrow}{A}$ U U G G C U C C A G C U $\overset{\downarrow}{U}$ U G C $\overset{\downarrow}{A}$ U G C $\overset{\downarrow}{A}$ U G A
G A G $\overset{\downarrow}{A}$ U A G U A G G G C A G U C C G

c X31 influenza RNA 8†
10 20
3′ U C G U U U U C G U C C C A C U G U $\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$
30 40
$\overset{\downarrow}{U}$ U G U $\overset{\downarrow}{A}$ $\overset{\downarrow}{U}$ U A C C U A G G U U U G U G
50 60
A C A C A G U U C G A A A G U C C A U C
70
U A A C G A A A G A A A C C G U A C A

5′-labelled RNA

d X31 influenza RNA 8‡
10 20
5′ A G U A G A A A C A A G G G U G U $\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$
30 40
$\overset{\downarrow}{U}$ $\overset{\downarrow}{U}$ $\overset{\downarrow}{A}$ U C A U U $\overset{\downarrow}{A}$ A A U $\overset{\downarrow}{A}$ A G C U G A A
A C G A

Arrows indicate sites of cleavage. * Refs 13 and 16. † McCauley and Skehel, unpublished. ‡ Ref. 16; Skehel, unpublished.

prepared from mouse Krebs ascites tumour cells and human lymphoblastoid (Daudi) cells and their enzyme activities compared with that of a rabbit reticulocyte lysate. Using 3'-labelled influenza virus RNA 7, similar patterns of cleavage were obtained with all three cell extracts: the 2-5A-dependent cleavages seen using the reticulocyte lysate (Fig. 1*c*) at nucleotides 17, 36 and 64 were also clearly observed using the mouse and human cell extracts (Fig. 3*a*). Similarly, with influenza virus RNA 8 (Fig. 3*b*), cleavages at identical positions (nucleotides 19, 20, 21, 25 and 26), although to different extents at nucleotide 25, were obtained in the presence of 2-5A using extracts from cells of all three types.

We conclude from these results that in extracts from cells of several different types, the 2-5A-dependent ribonuclease specifically cleaves at UpN sequences yielding 3'-phosphorylated oligonucleotides which terminate in the sequence UpNp. With the RNA substrates examined the sequences UA and UU were preferred. For the rabbit reticulocyte lysates the most susceptible sites of cleavage were at UA sequences, particularly sequences rich in U (residues 32 and 25 in Fig. 1*a*, *d*; bold figures in Table 1*a*, *c*). However, the preference for sites in U-rich regions was not obvious with the enzymes from the mouse and human cells. This emphasizes that although the 2-5A-dependent nucleases from different sources show similar specificities (Fig. 3), they may not be identical. Furthermore, the reticulocyte enzyme is known to be unusual in that it requires the tetramer (rather than the trimer) or higher oligomers of 2-5A for full activity[3].

Crude extracts were used in these experiments; thus it is possible that more than one nuclease is involved in any given extract. However, studies in progress on the purification of the nuclease and its specific interaction with 2-5A provide results against this.

Cleavage was not obtained at all UA and UU sequences. The reason for this is not known, but it is likely that the specificity is also influenced by the secondary structure of the substrate. In the parts of La Crosse sRNA and influenza RNA 7 examined, the only UA sequences at which cleavage was not observed were in regions of the molecule which may be double stranded[13,16]. However, the specificity of the nuclease reported does not simply reflect an unusual structure for the viral RNAs because the same specificity for UpN sequences was also observed with ribosomal RNA. Limited cleavages of rRNA have been observed in interferon-treated, SV40 virus-infected cells[17] and in cells into which 2-5A has been deliberately introduced[18]. Because specific rRNA cleavages can inactivate ribosome function (for example, those produced by colicin E3 in prokaryotes[19] or by reticulocyte RNase M[20]) it will be of considerable interest to determine whether rRNA cleavages are a common feature of interferon treatment and virus infection and whether

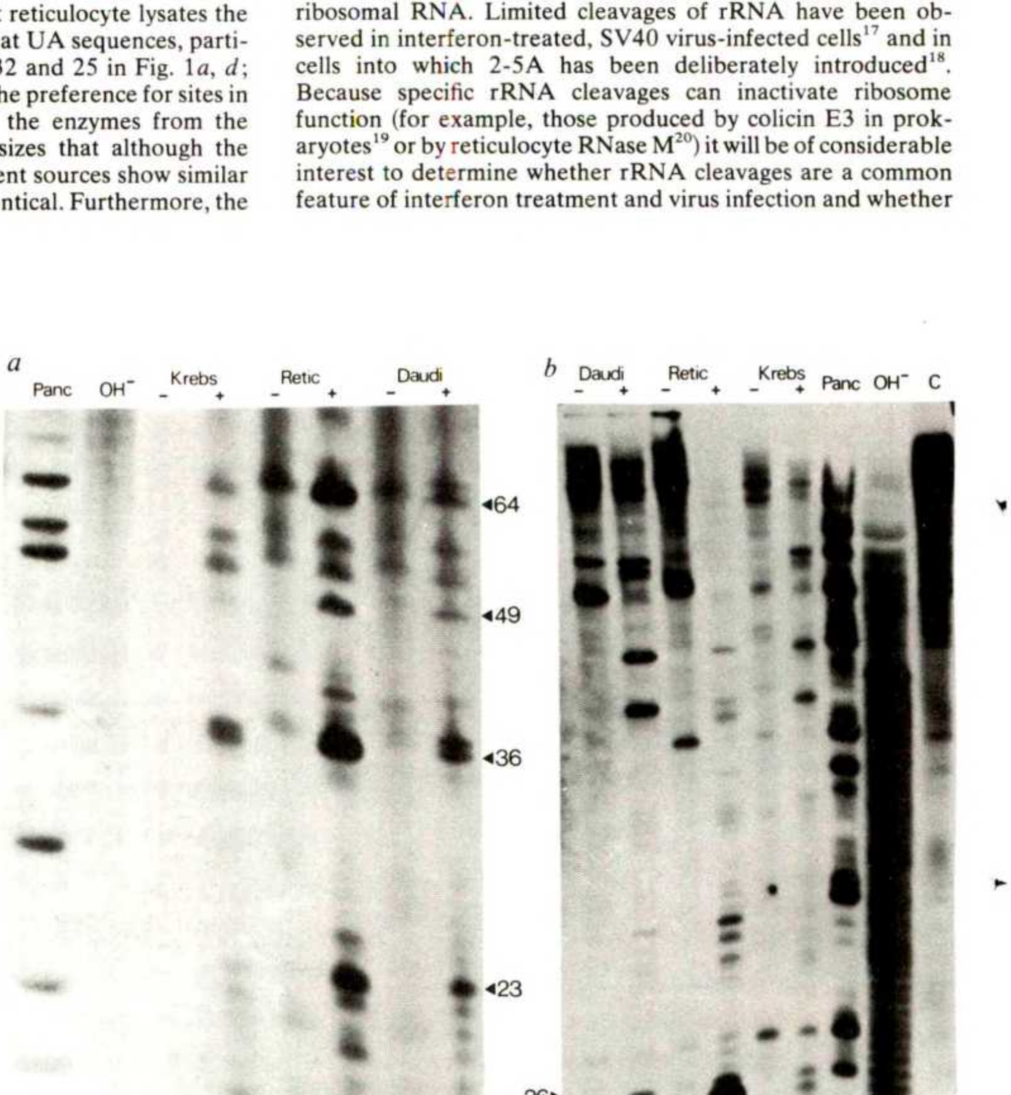


Fig. 3 Comparison of the specificities of the 2-5A-dependent ribonucleases prepared from rabbit reticulocytes, human Daudi lymphoblastoid and mouse Krebs tumour cells using 3'-labelled X31 influenza virus RNAs 7 and 8. 3'-Labelled influenza virus RNA 7(*a*) or RNA 8(*b*) were incubated for 60 min at 30 °C in a rabbit reticulocyte lysate (Retic), human Daudi lymphoblastoid extract (Daudi) or a mouse Krebs cell extract (Krebs) in the presence (+) or absence (−) of 100 nM tetramer 2-5A. Aliquots (10 μl) were diluted to 200 μl with 100 mM Tris *p*H 8.0, 100 mM NaCl, 1 mM EDTA and extracted twice with phenol/chloroform (1:1). The oligonucleotides were ethanol precipitated, redissolved in 7 M urea, 10 mM Tris *p*H 7.4 and applied to the polyacrylamide gel. The figure shows an autoradiogram of a 12% polyacrylamide gel (40×40×0.05 cm) containing the products of partial digestion with pancreatic RNase and partial alkaline hydrolysis, prepared as described in Fig. 2 legend. The numbers indicate the positions of cleavage by the 2-5A-dependent ribonucleases. Extracts of reticulocytes, Krebs and Daudi cells were prepared as described elsewhere[12,24], except that dialysis was omitted and the Duadi cells were homogenized in 40 mM KCl, 10 mM HEPES *p*H 7.6, 2.5 mM $MgCl_2$, 0.5 mM ATP and 2.5% glycerol. The Daudi and Krebs extracts were further centrifuged at 100,000*g* for 4 h to obtain the supernatant (S100) fraction. Incubations (50 μl) contained 20 μl of lysate (±20 μM haemin) or S100, and 1 mM ATP, 0.2 mM GTP, 10 mM creatine phosphate, 2 mM glucose, 75 mM KCl, 2 mM Mg acetate, 10 mM Tris-HCl buffer *p*H 7.6, a mixture of 19 amino acids (excluding cysteine), labelled RNA (50,000 c.p.m.) and tetramer 2-5A.

these or other nucleolytic cleavages in such cells show the same specificity as that reported here.

Since submission of this manuscript P. Lengyel and his colleagues (personal communication) have observed that of the four different ribohomopolymers poly(A), (U), (G) or (C), only poly(U) is cleaved by the 2-5A-dependent nuclease.

D. H. W. was the recipient of an EMBO fellowship. We thank Jack Obijeski for La Crosse virus RNA, Helen Donis-Keller for the *Physarum* RNase and Debora Peattie for communicating the chemical sequencing procedures.

Received 20 August; accepted 27 October 1980.

1. Kerr, I. M. & Brown, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 256–260 (1978).
2. Stark, G. R., Dower, W. J., Schimke, R. J., Brown, R. E. & Kerr, I. M. *Nature* **278,** 471–473 (1979).
3. Williams, B. R. G., Golgher, R. R., Brown, R. E., Gilbert, C. S. & Kerr, I. M. *Nature* **282,** 582–586 (1979).
4. Hovanessian, A. G., Wood, J., Meurs, E. & Montagniers, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 326–326S (1979).
5. Williams, B. R. G., Golgher, R. R. & Kerr, I. M. *FEBS Lett.* **105,** 47–52 (1979).
6. Clemens, M. J. & Williams, B. R. G. *Cell* **13,** 565–572 (1978).
7. Baglioni, C., Minks, M. A. & Maroney, P. A. *Nature* **273,** 684–687 (1978).
8. Williams, B. R. G., Kerr, I. M., Gilbert, C. S., White, C. N. & Ball, L. A. *Eur. J. Biochem.* **92,** 455–462 (1978).
9. Slattery, E., Ghosh, N., Samanta, H. & Lengyel, P. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4778–4782 (1979).
10. Schmidt, A. *et al. FEBS Lett.* **95,** 257–264 (1978).
11. Ball, L. A. & White, C. N. *Virology* **93,** 348–356 (1979).
12. Williams, B. R. G., Gilbert, C. S. & Kerr, I. M. *Nucleic Acids Res.* **6,** 1335–1350 (1979).
13. Obijeski, J. F., McCauley, J. & Skehel, J. J. *Nucleic Acids Res.* **8,** 2431–2438 (1980).
14. Scholtissek, C. *Curr. Topics Microbiol. Immun.* **80,** 139–169 (1978).
15. Peattie, D. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1760–1764 (1979).
16. Skehel, J. J. & Hay, A. J. *Nucleic Acids Res.* **5,** 1207–1219 (1978).
17. Revel, M., Kimchi, A., Schmidt, A., Shulman, L. & Chernajovsky in *Regulation of Macromolecular Synthesis* (eds Koch, G. & Richter, D) 284–302 (Academic, New York, 1979).
18. Hovanessian, A. G. & Wood, J. N. *Virology* **101,** 81–91 (1980).
19. Bowman, C. B., Sidikaro, J. & Nomura, M. *Nature new Biol.* **234,** 133–137 (1971).
20. Wreschner, D., Melloul, D. & Hertzberg, M. *Eur. J. Biochem.* **89,** 341–352 (1978).
21. England, T. E. & Uhlenbeck, O. C. *Biochemistry* **17,** 2069–2076 (1978).
22. Donis-Keller, H., Maxam, A. & Gilbert, W. *Nucleic Acids Res.* **4,** 2527–2538 (1977).
23. Rommelaere, J., Donis-Keller, H. & Hopkins, N. *Cell* **16,** 43–50 (1979).
24. Kerr, I. M., Brown, R. E., Clemens, M. J. & Gilbert, C. S. *Eur. J. Biochem.* **69,** 551–561 (1976).

Removal of O^6-methylguanine from DNA of normal and xeroderma pigmentosum-derived lymphoblastoid lines

Robert Sklar & Bernard Strauss*

Department of Microbiology, The University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

The ability to excise (repair) UV-induced pyrimidine dimers in *Escherichia coli* is not related to its ability to remove *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG)-induced O^6-methylguanine (O^6-MeG) from DNA[1]. It was therefore surprising that certain xeroderma pigmentosum cell lines, deficient in dimer excision, were also unable to remove O^6-MeG[1–3]. We find that removal of O^6-MeG occurs rapidly with a half life of less than 1 h. Two cell types can be distinguished: mex$^+$, which remove O^6-MeG residues produced by incubation with 0.5 μg ml^{-1} MNNG, and mex$^-$ cells, which are unable to remove the adduct. Xeroderma pigmentosum-derived lymphoblastoid lines of complementation groups A, C or D may be either mex$^+$ or mex$^-$. The biochemical mechanism for the removal of O^6-MeG in human cells is distinct from the excision of adducts produced by compounds such as *N*-acetoxy-*N*-2-acetylaminofluorene (AAAF) or by UV irradiation but it is not clear whether the distinction between mex$^+$ and mex$^-$ lines is genetic or epigenetic.

Treatment of cells or DNA with MNNG produces a variety of adducts in the DNA[4,5]. Of these, O^6-MeG is particularly important because it seems to lead to increased mutagenesis in prokaryotes and has been implicated as a potentially carcinogenic lesion[6–8]. Similar importance has been ascribed to the O^6-ethylguanine adducts produced by certain ethylating agents.

Cells have evolved systems for dealing with the O^6-alkylguanine lesion although the detailed biochemistry of the reaction remains to be determined[9,10]. The removal systems are unequally distributed in mammalian tissues: liver is efficient in removing O^6-alkylguanine compared with brain tissue which is inefficient[11]. It has also been observed that simian virus 40 (SV40)-transformed fibroblasts from patients with xeroderma pigmentosum are deficient in the ability to remove O^6-alkylguanine[2,3] and this observation has been extended to lymphoblastoid cells derived from an individual with xeroderma pigmentosum[1]. To determine whether the inability to remove O^6-MeG was genetically associated with the xeroderma characteristic, we decided to survey a variety of lymphoblastoid lines, including xeroderma heterozygotes, and determine whether they could remove the O^6-methyl adduct.

We first determined the kinetics of O^6-MeG and 3-methyladenine (3-MeA) removal from the DNA of Raji human

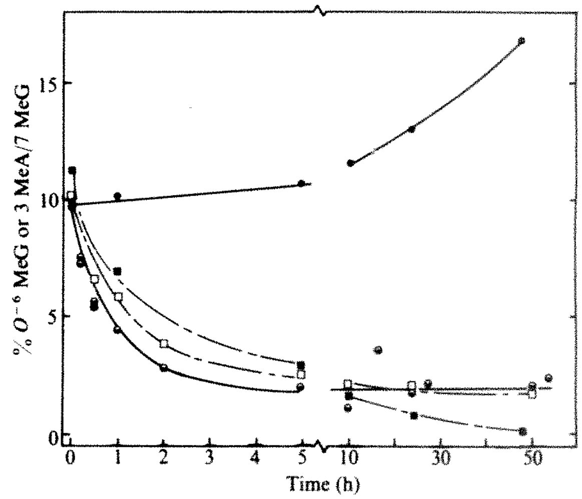


Fig. 1 Kinetics of O^6-MeG and 3-MeA removal in Raji and H2BT. Two litres of cells (~8 × 10^5 cells ml^{-1}) grown in spinner cultures were centrifuged and resuspended in 90 ml (Raji) or 60 ml (H2BT) of prewarmed SSC (0.15 M NaCl + 0.015 M sodium citrate) at 37 °C. Aliquots (10 ml) were treated for 2 min at 37 °C with 0.5 μg ml^{-1} ^{3}H-MNNG (112 mCi mmol^{-1}, Radiochemicals, Amersham) and 1 ml *N*-acetyl-L-cysteine (100 mM, *p*H 7.2) (ref. 12). Immediately after treatment a 0-time aliquot was lysed with 1 ml 2% SDS in SSC and frozen; later samples were centrifuged and resuspended in 100 ml RPMI 1640 plus 10% fetal calf serum and 4 mM L-glutamine in a 75 cm^2 Falcon flask and incubated with mild agitation for the time indicated. Cells were then centrifuged, resuspended in SSC and lysed with SDS as above. Cell lysates (10 ml) were treated for 1 h with 0.5 ml of 20 mg ml^{-1} heat-treated RNase (Worthington Biochemical) and for 4 h with 1.5 ml of 20 mg ml^{-1} heat-treated pronase (Calbiochem-Behring). Digested lysates were extracted three times with redistilled phenol saturated with SSC and neutralized with 0.2 volumes of 1 M Tris base. DNA was precipitated by adding 0.1 volume of 2.5 M sodium acetate and 2 volumes of 2-ethoxyethanol, washed twice with ice-cold 100% ethanol and dried. The DNA was dissolved in 100 μl of 0.05 M ammonium formate with 0.02% sodium azide, *p*H 7.5, and hydrolysed at 70 °C for 20 min after adding 0.2 volumes of 1 M HCl and a 50 μl mixture of 3-MeA (2 mg ml^{-1}, Cyclo Chemical), 7-MeG (5 mg ml^{-1}, Sigma), and O^6-MeG (5 mg ml^{-1}) (O^6-MeG, obtained from Dr D. Ludlum, was hydrolysed to yield the free base). The hydrolysate was spotted over 6 cm of a 0.5-mm cellulose MN300 thin-layer chromatography plate pre-chromatographed with solvent (Anal-tech) and developed in 2-propanol : 28% NH_4OH : H_2O (7 : 2 : 1) (ref. 11). UV absorbing bands corresponding to the above markers and guanine were removed from the plate with a razor blade and placed in scintillation vials, then 2 ml of 0.05 M ammonium formate, 0.02% sodium azide with 0.1 volume of 1 M HCl were added. Aquasol was added to the radioactive bands and the vials were vortexed and counted after several hours of equilibration. The guanine band was centrifuged twice in a microcentrifuge to remove the cellulose and diluted for absorbancy reading at 260 nm. The amount of DNA (in μg) was determined by multiplying the total guanine absorbancy by 164, a number determined from the absorbancy of a known amount of guanine and the molar ratio of guanine in DNA. Squares represent ratio of 3-MeA to 7-MeG × 100; circles, ratio of O^6-MeG to 7-MeG × 100; completely filled symbols, line H2BT; open squares and half-filled circles, lymphoma line Raji.

---

* To whom correspondence should be addressed.

0028-0836/81/050417—04$01.00

lymphoma cells because this line is known to remove methylation products from DNA[1]. Exponentially growing cells were treated with a pulse of MNNG and at intervals portions of the culture were collected and the alkylation products analysed. Figure 1 shows that O^6-MeG was rapidly removed from the DNA with a half life of less than 1 h at 37 °C, and that at the concentration of MNNG used (0.5 μg ml^{-1}) the ratios of O^6-MeG and 3-MeA to 7-methylguanine (7-MeG) declined at about the same rate. Experiments at higher MNNG concentrations indicate that the capacity for O^6-MeG removal is much more limited than for 3-MeA. The results in Fig. 1 are expressed as a ratio of radioactivity in O^6-MeG to 7-MeG and because 7-MeG is gradually lost from the DNA[1], the inability to remove the O^6-MeG adduct from the DNA results in an increase in this ratio with time of incubation in strain H2BT.

The result that the H2BT line, derived from a spherocytosis patient and heterozygous for thymidine kinase[13], was deficient in O^6-MeG removal was unexpected, as was the discovery that other non-xeroderma lines were also deficient (Table 1). In a series of control experiments in the Raji line at different MNNG doses we found that the 0-time value for the O^6-MeG/7-MeG ratio was 10.4% and the 0-time 3-MeA/7-MeG ratio was 10.1%. All the lines which remove O^6-MeG have O^6-MeG/7-MeG ratios which are consistent with the data obtained with

Table 1 Removal of O^6-MeG by human lymphoblastoid lines

Strain Designation	Origin		Incubation time (h)	7-MeG (d.p.m. per μg DNA)	O^6-MeG/ 7-MeG × 100	3-MeA/ 7-MeG × 100	Classification
a Burkitt's lymphoma lines							
Raji	Burkitt's lymphoma		0 27.5	13.52 3.59	10.9 1.2	9.76 2.15	mex$^+$
Daudi	Burkitt's lymphoma		0.5 10.0	2.04 1.99	8.0 3.3	3.8 2.8	mex$^+$
HR-1	Burkitt's lymphoma		0 3.0	— 4.46	7.1 0.73	10.2 2.5	mex$^+$
b Human lymphoblastoid lines							
H2BT	Thymidine kinase heterozygote derived from a spherocytosis patient		0 24.0	10.85 2.93	9.7 13.0	11.3 1.2	mex$^-$
L 33-6-1	Infectious mononucleosis		0 24.3	7.56 2.03	10.2 15.4	9.47 0.4	mex$^-$
GM 892	Normal female (E. Conod, 1970) from HGMCR*		0 24.0	8.34 4.43	10.0 12.9	9.14 1.4	mex$^-$
GM 621	Normal male established 6/4/74 from HGMCR		0 24.0	9.21 4.42	9.83 11.7	9.85 1.5	mex$^-$
GM 1953	Normal female established 3/17/75 from HGMCR		0 3.0	— —	8.09 3.05	6.62 3.25	mex$^+$
D 8-5-1	Progeria patient†		0.5 29.0	9.05 4.18	3.94 0.82	7.86 1.54	mex$^+$
D 16-1-1	Werner's Syndrome†		0 25.5	11.63 8.43	9.20 2.27	11.4 1.31	mex$^+$
c Xeroderma pigmentosum-related lines							
		Complementation group					
GM 2250	HGMCR‡	A	0 24.0	3.1 1.7	8.3 1.5	11.7 2.2	mex$^+$
GM 2500	HGMCR‡	A	0 24.0	11.5 11.6	10.0 2.5	— —	mex$^+$
GM 2246	HGMCR‡	C	0 24.0	12.7 7.6	9.4 1.6	11.5 1.3	mex$^+$
GM 2248	HGMCR‡	C	0 24.0	— —	10.5 11.8	11.3 1.4	mex$^-$
GM 2498	HGMCR‡	C	0 24.0	8.1 5.1	10.4 14.0	11.6 1.6	mex$^-$
GM 2253	HGMCR‡	D	0.5 3.0	— —	3.81 2.02	5.82 2.72	mex$^+$
GM 2473	HGMCR‡	D heterozygote mother of 2253	0.5 3.0	— —	9.78 9.86	6.79 3.88	mex$^-$
D 70-1-2	Xeroderma heterozygote (father, family JAGR)†	Unknown	0 3.0	8.74 9.45	6.85 1.6	9.68 2.7	mex$^+$
D 70-2-2	Xeroderma heterozygote (mother, family JAGR)†	Unknown	0 23.0	— —	9.9 3.2	10.5 1.7	mex$^+$
D 70-3-1	Xeroderma proband, family JAGR†	Unknown	0 3.0	4.25 4.67	8.44 1.7	8.44 3.8	mex$^+$
D 70-3-2	Xeroderma proband†. Independent lymphoblastoid culture.	Unknown	0 32.0	5.15 2.92	11.9 2.5	14.8 1.8	mex$^+$
A 64-1-2	Xeroderma heterozygote (mother), MALO family.†	Unknown	0.5 29.0	14.4 6.06	3.41 0.38	8.14 1.0	mex$^+$
A 64-2-3	Xeroderma proband, male, MALO family.†	Unknown	0 24.0	8.25 4.56	7.74 0.93	9.27 1.2	mex$^+$
D 75-1-2	Xeroderma, male, independent culture from A 64-2-3.†	Unknown	0 25.0	6.19 4.0	8.79 3.0	10.7 1.6	mex$^+$
D 75-2-1	Xeroderma patient, MALO family, sister of A 64-2-3.†	Unknown	0 25.0	11.75 8.67	8.68 3.0	9.71 1.3	mex$^+$
B 5-3-1	Xeroderma heterozygote, MALO family father.†	Unknown	0 25.0	11.48 4.94	9.25 3.23	10.3 1.5	mex$^+$

500 ml of cells (~8×10^5 ml^{-1}) grown in spinner culture were collected by centrifugation and resuspended in 20 ml SSC. Aliquots (10 ml) were treated with 0.5 μg ml^{-1} ^{3}H-MNNG (112 mCi mmol^{-1}), resuspended in RPMI 1640 as described in Fig. 1. legend and incubated for the indicated time. The cells were then collected and lysed and the DNA purified as described in Fig. 1 legend. The Burkitt's lymphoma lines were obtained from P. Gerber (Raji) and G. Klein (Daudi and HR-1). H2BT supplied by W. Thilly; L 33-6-1 from J. Littlefield.

* HGMCR, Human Genetic Mutant Cell Repository, Camden, New Jersey.

† From Earl Henderson.

‡ From A. Andrews.

Table 2 Excision repair in mex$^+$ and in a mex$^-$ lymphoblastoid line

			Repair synthesis (c.p.m. μg^{-1}) induced by:										
			AAAF (μg ml^{-1})		AAAF/ control	MMS (μg ml^{-1})			MMS/ control	MNNG (μg ml^{-1})			MNNG/ control
Strain	Origin	Control	12.5	25	(25 μg)	110		220	(440 μg)	1	2.5	5	(5 μg)
D 70-1-2	JAGR father (mex$^+$)	13.2		128	(9.7)								
D 70-2-2	JAGR mother (mex$^+$)	11.1		138	(12.4)								
D 70-3-1	JAGR proband (mex$^+$)	16.6		22	(1.3)								
D 70-3-2	JAGR proband (mex$^+$) (2nd culture)	12.9		22	(1.7)								
D 70-4-1	JAGR brother	28.2		170	(6.0)								
D 70-5-2	JAGR sister	18.3		136	(7.4)								
D 70-1-2	JAGR father (mex$^+$)	79.0	476	733	(9.3)	192	294	285	(3.6)	129	272	424	(5.4)
GM 621	Normal male (mex$^-$)	37.0	286	398	(10.8)	162	220	179	(4.8)	88	169	243	(6.6)

DNA excision repair was measured by the BND cellulose method[16]. Rapidly growing cells were diluted to 4×10^5 ml^{-1}, grown overnight, concentrated to about 10^6 ml^{-1}, and 2 ml of suspension used for determination. Repair synthesis was measured in the presence of 2 mM hydroxyurea and 10^{-6} M fluorodeoxyuridine along with 5 μCi ml^{-1} ^{3}H-thymidine (15 Ci mmol^{-1}). Results were not corrected for pool size and the ratio of drug-induced counts to the background is given as an indication of the relative repair capacity. Determinations of repair induced by MMS, AAAF and MNNG were done separately from the tests of the JAGR family, which were carried out by Mr Dino Vallera.

Raji (Fig. 1). Of 26 lymphoma and lymphoblastoid lines classified for their ability to remove O^6-MeG, we found 19 removers and 7 non-removers. We call the non-removers mex$^-$ (methyl excision minus) for convenience but the terminology has no genetic significance.

Mex$^+$ strains include three Burkitt's lymphoma lines, lines derived from xeroderma patients of complementation groups A, C and D, and representatives of tx xeroderma families of unknown complementation group. All xeroderma lines were tested in this laboratory for repair synthesis induced by AAAF and found to be repair negative compared with controls including the xeroderma heterozygotes. As the lymphoblastoid cultures are continuous lines, this result indicates that transformation does not affect repair deficiency. The seven mex$^-$ lines we have identified include a line derived from an infectious mononucleosis patient, a spherocytosis patient, a xeroderma heterozygote and a normal male and female culture.

The dose of MNNG used in these experiments (0.5 μg ml^{-1}) reduced the survival of mex$^+$ Raji cells to ~20%, as determined in a cloning experiment. Inability to remove O^6-MeG adducts after treatment at 0.5 μg ml^{-1} may indicate sensitivity of MNNG at lower concentrations for mex$^-$ strains. Most lymphoblastoid lines cannot be efficiently cloned in soft agar but their growth can be followed by cell counts in liquid medium[14]. We have observed (Fig. 2) that mex$^-$ cells are more sensitive to growth inhibition at low MNNG doses than are removal-competent strains (mex$^+$). The greater sensitivity of mex$^-$ strains was not found when treated with the carcinogen AAAF whose adducts are removed by UV-like (nucleotide) excision[15]. The mex$^-$ strain GM 892 was also more sensitive to growth inhibition by 3 or 5 μg ml^{-1} *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG) than the mex$^+$ D70-1-2, and ENNG pretreatment (2.5 μg ml^{-1}) 2 h before MNNG challenge eliminated the capacity of the mex$^+$ Raji line to remove O^6-MeG, suggesting that the mex system is involved in the metabolism of ethyl adducts.

Both mex$^+$ and mex$^-$ strains respond to treatment with methyl methanesulphonate (MMS), MNNG or AAAF with increased DNA excision-repair synthesis (Table 2) and estimation of the relative activities by comparison of the ratio of repair synthesis with the controls (Table 2) indicates that the mex$^+$ D70-1-2 and mex$^-$ GM 621 give an almost equivalent repair synthesis response. Inability to respond to AAAF by increased repair synthesis is not related to the ability to remove O^6-MeG in lymphoblastoid cells (Tables 1 and 2).

The result that mex$^-$ strains respond to treatment with MNNG by increased repair synthesis seems to indicate that MNNG produces a variety of lesions, some of which (for example, 3-MeA) are removed, producing apurinic sites which can be repaired by an intact AP (apurinic/apyrimidinic) repair system, and either that O^6-MeG removal is a minor reaction which does not show up in the repair measurements at the high MNNG dose used, or that removal of the O^6-methyl group is not accompanied by base removal as suggested by *in vitro* studies[9,10].

We do not know the genetic basis for the mex$^-$ characteristic. There could be a major mex polymorphism in human populations and this hypothesis could account for a mex$^+$ xeroderma line and a mex$^-$ heterozygous line from the same family (Table 1). Alternatively, some epigenetic process might occur in cultured cells resulting in a mex$^+$ to mex$^-$ conversion. Such processes are known to occur in normal development because different organs of the same animal (for example, brain and liver) can have a different mex characteristic[11]. If the process is epigenetic, it is possible that functions other than the ability to remove O^6-MeG would be simultaneously affected.

Our mex$^+$ and mex$^-$ strains have properties like those studied by Day *et al*[17,18] who reported strains which differ in their ability to reactivate MNNG-treated adenovirus. We have found that a pair of Day's mer$^-$ and mer$^+$ are mex$^-$ and mex$^+$ respectively and this observation has been confirmed (Day, personal communication). Separation of the mex characteristic from xeroderma phenotype also agrees with the finding that only a SV40-transformed xeroderma cell line was hypersensitive to sister chromatid induction by the alkylating agent, ethyl methanesulphonate[19].

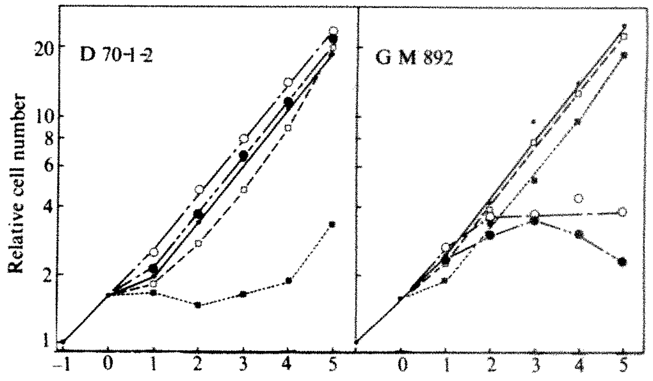


Fig. 2 Effect of MNNG and AAAF on the growth of the mex$^+$ (D70-1-2) and mex$^-$ (GM 892) lymphoblastoid lines. Cultures were maintained in a rapidly growing state by daily dilution to 4×10^5 ml^{-1} in RPMI 1640+16% fetal calf serum. At day '0', the cells were diluted to 4×10^5 ml^{-1} and treated with AAAF or MNNG for 1 h at 37 °C. The cells were then collected by centrifugation, resuspended in medium at 4×10^5 ml^{-1} and incubated at 37 °C. Cultures were counted daily using a Coulter counter and the population density was adjusted to 4×10^5 ml^{-1}. The daily increase in population size was used to calculate a relative cell number assuming no dilution. AAAF was dissolved in dimethyl sulphoxide, which was added in equal amounts to all cultures. Small •, control; large ○, MNNG (0.1 μg ml^{-1}); large ●, MNNG (0.2 μg ml^{-1}); □, AAAF (4 μg ml^{-1}); ■, AAAF (8 μg ml^{-1}).

We thank Dr Earl Henderson for supplying mex$^+$ strains. This work was supported by NCI contract NO1-CP-85669, and by grants from the NCI (CA 14599) and the Department of Energy (EY75 S-02 2040). R.S. was a trainee on a NCI-sponsored training grant in environmental biology (CA 09273).

Received 7 August; accepted 7 November 1980.

1. Altamirano-Dimas, M., Sklar, R. & Strauss, B. *Mutat. Res.* **60**, 197–206 (1979).
2. Goth-Goldstein, R. *Nature* **267**, 81–82 (1977).
3. Bodell, W., Singer, B. & Cleaver, J. *Nucleic Acids Res.* **6**, 2819–2829 (1979).
4. Warren, W. & Lawley, P. *Carcinogenesis* **1**, 67–78 (1980).
5. Singer, B. *Nature* **264**, 333–339 (1976).
6. Loveless, A. *Nature* **223**, 206–207 (1969).
7. Schendel, P. & Robins, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6017–6020 (1978).
8. Sklar, R. & Strauss, B. *J. molec. Biol.* **143**, 343–362 (1980).
9. Karran, P., Lindahl, T. & Griffin, G. *Nature* **280**, 76–77 (1979).
10. Pegg, A. *Biochem. biophys. Res. Commun.* **84**, 166–173 (1978).
11. Goth, R. & Rajewsky, M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 639–643 (1974).
12. Lawley, P. & Thatcher, C. *Biochem. J.* **116**, 693–707 (1970).
13. Skopek, T., Liber, H., Penman, B. & Thilly, W. *Biochem. biophys. Res. Commun.* **84**, 411–416 (1978).
14. Higgins, N. P. & Strauss, B. *Cancer Res.* **39**, 312–320 (1979).
15. Regan, J. & Setlow, R. *Cancer Res.* **34**, 3318–3325 (1974).
16. Scudiero, D., Henderson, E., Norin, A. & Strauss, B. *Mutat. Res.* **29**, 473–488 (1975).
17. Day, R. & Ziolkowski, C. *Nature* **279**, 797–799 (1979).
18. Day, R., Ziolkowski, C., Scudiero, D., Meyer, S. & Mottern, M. *Carcinogenesis* **1**, 21–32 (1980).
19. Heddle, J. & Arlett, C. *Mutat. Res.* **72**, 119–125 (1980).

Assembly units of clathrin coats

Ernst Ungewickell & Daniel Branton

The Biological Laboratories, 16 Divinity Avenue,
Harvard University, Cambridge, Massachusetts 02138

Clathrin, a polypeptide of molecular weight (MW) 180,000, is the main constituent of the polygonal network that forms the coat of coated pits and vesicles[1]; these vesicles play a part in intracellular transport between membranous organelles[2,3]. This function involves specific recognition of target membranes as well as fusion and fission events that must be coordinated with the assembly, partial disassembly or reorganization of the clathrin coats. To understand these interactions on a molecular level, information about the structure of clathrin and the interactions of clathrin with itself and other proteins is required. Here we show that purified clathrin coats dissociate reversibly into triskelions, structures composed of three usually bent, rather flexible legs radiating from a centre. We have determined the molecular weight of these triskelions and conclude that they contain trimers of clathrin together with about three light molecular weight polypeptide chains.

Coated vesicles were isolated from fresh calf brains according to the procedure of Keen *et al.*[4] and the clathrin coats were dissociated in 2 M urea[5]. Two successive gel filtrations separated clathrin from polypeptides with apparent molecular weights of ~90,000–125,000 and 50,000–60,000, but failed to remove two light chains with apparent molecular weights of ~36,000 and 33,000 (Fig. 1). Experiments in which this column-purified clathrin was reacted with the cleavable cross-linking reagent 3,3′-dithiobis-propionimidate confirmed Pearse's[6] suggestion that these two light chains are bound to clathrin (results not shown). From the relative intensities of the Coomassie Brilliant Blue staining, we estimated that there is approximately one light chain per clathrin polypeptide. This ratio did not vary significantly between preparations although we did observe that the clathrin light chains are very sensitive to proteolysis.

It has been reported that antibodies to two polypeptides in the 30,000–40,000 MW range (probably the light chains) cross-react with brain tropomyosin[7]. We compared the electrophoretic mobilities of brain tropomyosin and clathrin light chains in SDS-polyacrylamide gels and in SDS-polyacrylamide gels containing urea. Tropomyosins from skeletal muscle and calf brain show anomalously low electrophoretic mobility in urea/SDS-polyacrylamide gels[8]. In these conditions brain tropomyosin with an apparent molecular weight of 30,000 in SDS

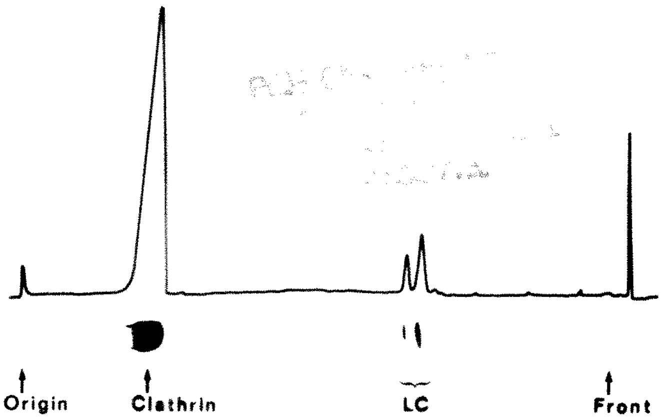


Fig. 1 Polyacrylamide gel electrophoresis of purified dissociate clathrin coats. Coated vesicles were prepared according to the procedure of Keen *et al.*[4] except that all solutions contained 0.005% phenylmethylsulphonylfluoride (PMSF). To dissociate the coats, purified coated vesicles were dialysed overnight at 4 °C against 500 volumes of 30 mM Tris-HCl, 20 mM NaCl, 1 mM EDTA, 0.2 mM dithiothreitol (DTT), 2 M urea, 0.02% NaN_3 and 0.005% PMSF, *p*H 7.5. Vesicles were removed by centrifugation at 100,000*g* for 1 h. The supernatant, which contained large amounts of protein in addition to clathrin, was chromotographed at 4 °C on a Sepharose 4B-Cl gel filtration column (90×2.5 cm), which was equilibrated with dialysis buffer. Clathrin eluted together with two low-MW polypeptides (light chains) and small amounts of polypeptides with apparent molecular weights of ~100,000 and 55,000 (molecular weights were estimated from the relative mobilities of polypeptides in SDS-polyacrylamide gels) which were completely removed by a second gel filtration in triethanolamine. The clathrin-containing fractions from the first column were concentrated by precipitation with 50% ammonium sulphate and immediately dissolved in 50 mM triethanolamine, 20 mM NaCl, 1 mM EDTA, 0.2 mM DTT and 0.02% NaN_3, *p*H 8.3. After dialysis against the same buffer, the protein was chromatographed in the same medium at 4 °C on a Sephadex 4B-Cl gel filtration column (100×1.5 cm). The polypeptide composition of the final pooled clathrin-containing fractions was analysed in SDS-polyacrylamide gradient gels (6–15%) using the discontinuous buffer system of Laemmli[9]. The gel was scanned with a Joyce Loebl microdensitometer. LC, clathrin light chains.

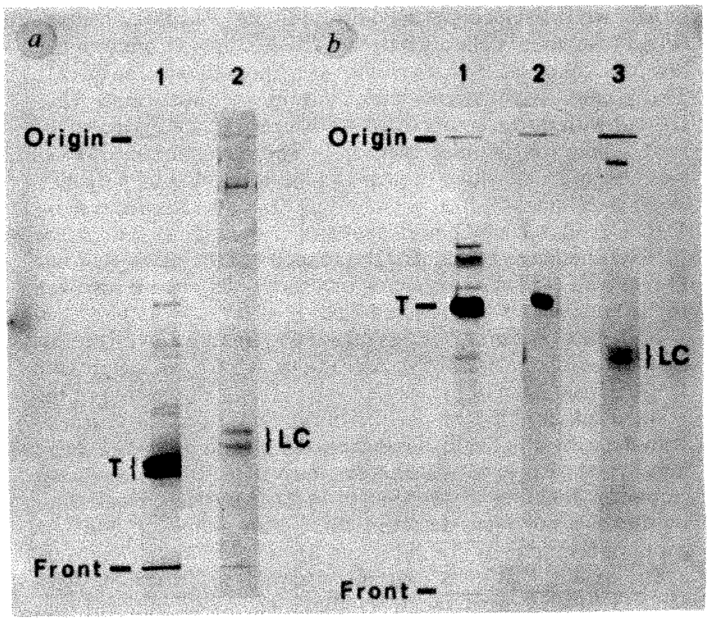


Fig. 2 Comparison between the electrophoretic behaviour of brain tropomyosin and clathrin light chains in SDS- and urea/SDS-polyacrylamide gels. *a*, 10% SDS-polyacrylamide gel[8]: lane 1, brain tropomyosin (T); lane 2, clathrin light chains (LC). *b*, urea/SDS-polyacrylamide gel: lane 1, brain tropomyosin; lane 2, skeletal muscle actin; lane 3, clathrin light chains.

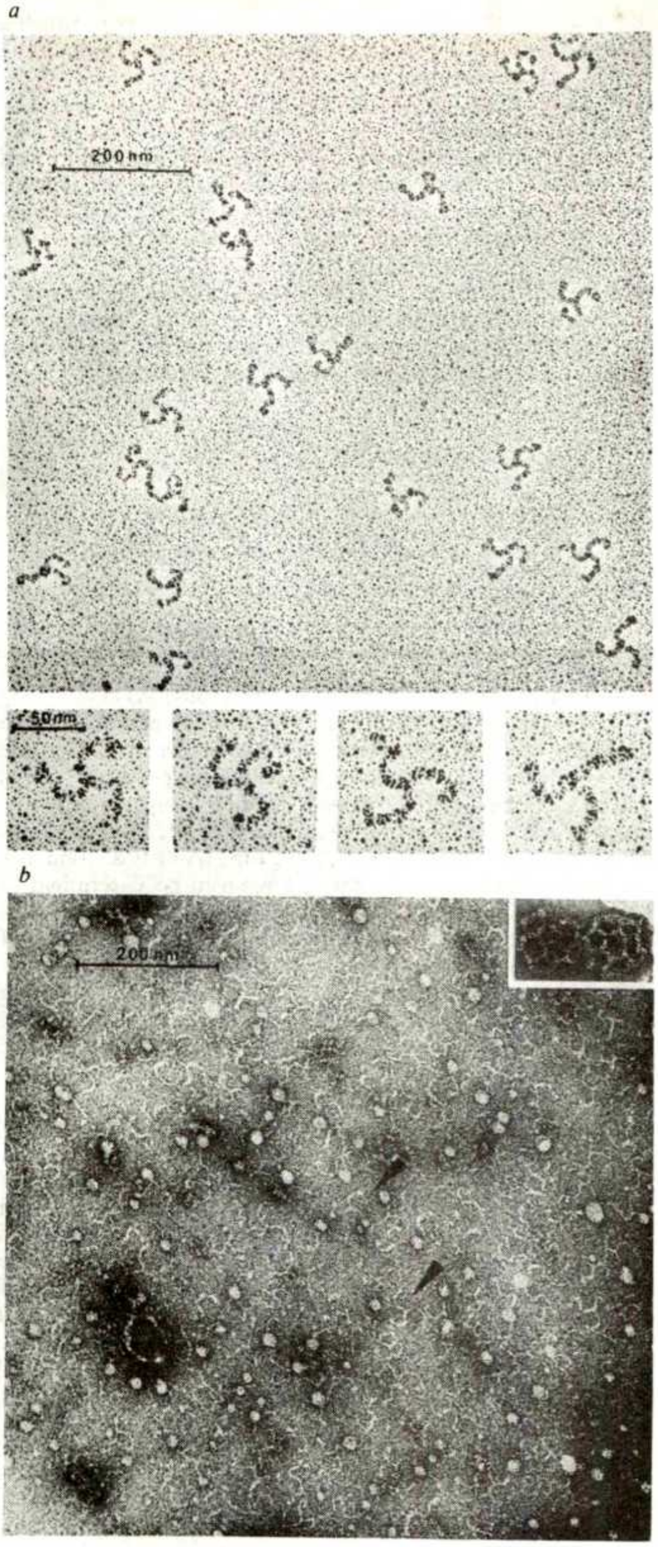


Fig. 3 Electron micrographs of 8.4 *s* clathrin triskelions, after *a*, rotary shadowing and *b*, negative staining with uranyl acetate. For rotary shadowing, dissociated clathrin coats (20 μg ml^{-1}) in 50 mM triethanolamine, *p*H 8.3, were mixed at room temperature with two volumes of glycerol and sprayed onto freshly cleaved mica. The molecules were then dried at room temperature under vacuum in a Balzers freeze-etch device, rotary-shadowed at an angle of 4.5° with Pt/C and coated from above with pure carbon (for details of method see ref. 11). For negative staining dissociated clathrin coats in 0.3 mM Tris-HCl, *pH* 8.0 (20 μg ml^{-1}) were applied to Formvar- and carbon-coated grids and negatively stained with 1% uranyl acetate. Arrows in *b* mark structures which resemble those seen in *a*. Inset in *b* shows coats reassembled from purified triskelions after dialysis against buffer A. All specimens were examined in a Philips EM 301 electron microscope at 80 kV. *a*, ×90,000 and 180,000; *b*, ×90,000.

gels has almost the same electrophoretic mobility as actin (43,000 MW) (ref. 8 and Fig. 2). In contrast, the clathrin light chains which ran with apparent molecular weights of 36,000 and 33,000 in SDS did not show a low electrophoretic mobility in SDS gels containing urea (Fig. 2). From this observation and from results of one-dimensional peptide maps (not shown) we conclude that the light chains and brain tropomyosin are not closely related proteins.

Analytical ultracentrifugation showed that the purified clathrin–light chain complex had a sedimentation coefficient, $s_{20,w}$, of 8.4 in 50 mM triethanolamine, 20 mM sodium chloride, 1 mM EDTA at *p*H 8.3. On dialysis against 0.1 M 2-(*N*-morpholino)-ethanesulphonic acid, 0.5 mM magnesium chloride, 0.2 mM EGTA and 0.002% sodium azide, *p*H 6.5 (buffer A), the 8.4 *s* material formed typical coats or baskets with relatively uniform diameters of ~800 Å (Fig. 3*b* inset). In several reassembly experiments, 75–93% of the total protein was recovered in the pelleted baskets after centrifugation at 100,000*g* for 1 h. The polypeptide composition of the pellets was identical to that of the starting material.

After rotary shadowing or negative staining, the 8.4 *s* material appeared as triskelions with a mean leg length of 445 Å ± 23 (s.d.). The legs in most of the molecules showed a preferred direction of curvature. When the curvature was sharp, it was located ~190 Å from the centre point where the three legs join. At higher magnification each leg seems to be composed of two strands which are twisted around each other and which seem to form a hairpin loop at the end.

The molecular weight of triskelions was determined by electron microscopy[12] and sedimentation equilibrium. For electron microscopy we mixed triskelions at a known weight concentration (determined by colorimetric micro-Kjeldahl analysis[13] using a nitrogen factor of 6.33 (ref. 14)) with myosin (at a known weight concentration), and counted the ratio of triskelions to myosin molecules in electron micrographs of droplets that had been dried and rotary shadowed. The estimated molecular weight was 627,000 ± 55,000 (s.d.; total number of particles counted was ~2,000). For sedimentation equilibrium the high-speed Yphantis method was used[15]. An apparent molecular weight of 630,000 was calculated from the plot in Fig. 4.

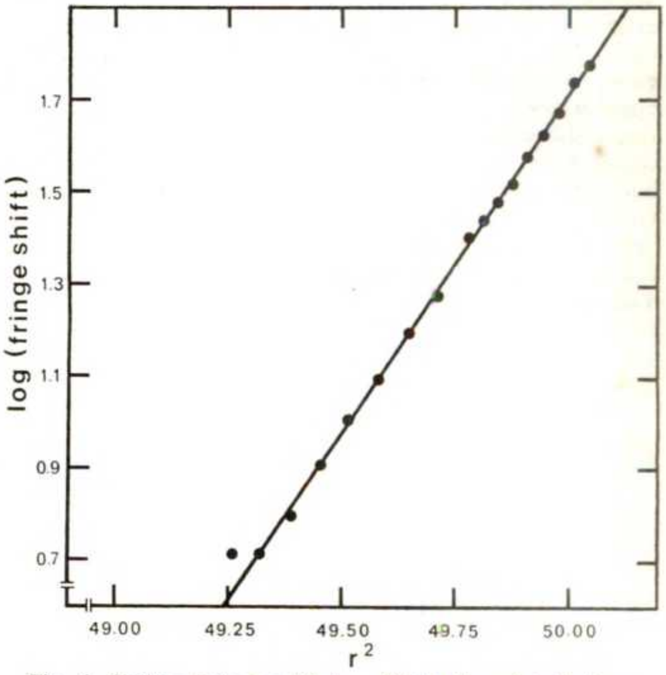


Fig. 4 Sedimentation equilibrium distribution of triskelions. Sedimentation equilibria were performed by the high speed method[15], with the use of the Rayleigh interference optical system in a Spinco model E ultracentrifuge at 9,341 r.p.m. and 9.1 °C. The protein concentration was 0.5 mg ml^{-1}, and the partial specific volume calculated from the amino acid composition was 0.75 cm^3 g^{-1} (ref. 14).

Assuming a 180,000 MW for a single clathrin polypeptide[1], the most straightforward interpretation of our data is that triskelions comprise three clathrin monomers and probably three light chains. We do not know whether a single triskelion contains both the 36,000 and 33,000 MW light chains. Because triskelions are stable molecular structures observed in various conditions, it is probable that other investigators studying the properties of dissociated clathrin coats have been dealing with triskelions (for example, the 8.8 *s* preparation of Puszkin and coworkers[16,17]).

According to Crowther *et al.*[18], a clathrin coat which is constructed of 8 hexagons and 12 pentagons (36 vertices) contains approximately 100 clathrin monomers. This suggests that each vertex in the coat is the centre of a triskelion. A triskelion leg is long enough to span two edges of the polygons that make up the coat[10,18] but other arrangements in which a portion of each leg could extend out of the coat towards the contained vesicle cannot be excluded.

We thank F. Marty, A. Verkleij, D. Caspar, S. Lehrer and D. Siegel for encouragement and discussions, W. Gratzer for help with the hydrodynamic measurements (which were performed at the MRC Cell Biophysics Unit in London), and Ken Giedd for technical assistance. E.U. was supported by a Training Grant from the Deutsche Forschungsgemeinschaft; the research was supported by NIH grant GM 20396.

Note added in proof: T. Kirchhausen and S. C. Harrison (Harvard) have recently shown by cross-linking experiments that the 8 *s* species contains three heavy chains, each in close contact with a light chain (manuscript submitted for publication).

Received 23 September; accepted 18 November 1980.

1. Pearse, B. M. F. *J. molec. Biol.* **97**, 93–98 (1975).
2. Goldstein, J. L., Anderson, R. G. W. & Brown, M. S. *Nature* **279**, 679–685 (1979).
3. Pearse, B. *Trends biochem. Sci.* **5**, 131–134 (1980).
4. Keen, J. H., Willingham, M. C. & Pastan, J. H. *Cell* **16**, 303–312 (1979).
5. Schook, W., Ores, C. & Puszkin, S. *J. Cell Biol.* **15**, 119a (1977).
6. Pearse, B. M. F. *J. molec. Biol.* **126**, 803–812 (1978).
7. Puszkin, S., Maimon, J. & Schook, W. *J. Cell Biol.* **83**, 293a (1979).
8. Bretscher, A. & Weber, U. *FEBS Lett.* **85**, 145–148 (1978).
9. Laemmli, U. U. *Nature* **227**, 680–685 (1970).
10. Heuser, J. *J. Cell Biol.* **84**, 560–583 (1980).
11. Tyler, J. M. & Branton, D. *J. ultrastruct. Res.* **71**, 95–102 (1980).
12. Williams, R. C. & Backus, J. *J. Am. chem. Soc.* **71**, 4052–4057 (1949).
13. Jaenicke, L. *Analyt. Biochem.* **61**, 623–627 (1974).
14. Pearse, B. M. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1255–1259 (1976).
15. Yphantis, D. A. *Biochemistry* **3**, 297–317 (1964).
16. Schook, W., Puszkin, S., Bloom, W., Ores, C. & Kochwa, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 116–120 (1979).
17. Bloom, W. S., Schook, W., Feageson, E., Ores, C. & Puskin, S. *Biochim. biophys. Acta* **598**, 447–455 (1980).
18. Crowther, R. A., Finch, J. T. & Pearse, B. M. F. *J. molec. Biol.* **103**, 785–798 (1976).

Disulphide bonds of haemagglutinin of Asian influenza virus

Michael Waterfield & Geoffrey Scrace

Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields,
London WC2A 3PX, UK

John Skehel

National Institute for Medical Research, Mill Hill,
London NW7 1AA, UK

The complete amino acid sequences of the haemagglutinins (HAs) of one strain of each of the Hav1 and H2 subtypes and of three strains of the H3 subtype of influenza virus have been established[1–3]. A large external fragment (BHA) of HA can be released from virus particles using the protease bromelain[4] that cleaves the protein close to its carboxyl terminus, which is inserted in the virus membrane (see Fig 1). Here we show that a single disulphide bond joins the two component polypeptide chains BHA_1 and BHA_2 and establish the location of five intrachain disulphides. These disulphide bonds have been located in the three-dimensional structure of HA which is known to 3 Å resolution[5,6].

Sequence analysis of HA of infectious A/Japan/305/57(H2N1) shows that the subunit contains 546 amino acids, 16 of which are half-cystine residues[3]. Analyses of BHA_1 and BHA_2 calculated as 324 and 170 residues respectively[7] reveal 9.1 and 2.9 half-cystine residues in these polypeptides. The precise location of the bromelain cleavage site(s) is unknown because the enzyme does not cleave the HA of this strain at a single site. However, cleavage is known to be C-terminal to half-cystine 473 in BHA_2 (ref. 8). Established sequence data suggest that four of the 16 half-cystines are in the C terminus of HA_2, which is removed by bromelain treatment. Alkylation of half-cystine residues of BHA in 6 M guanidine hydrochloride using ^{14}C-iodoacetamide failed to reveal any free sulphydryl groups, thus all the 12 half-cystines in BHA must exist in disulphide bonds.

To locate the half-cystines linked by disulphide bonds, fragments containing cystine residues were purified. BHA was first cleaved with cyanogen bromide and fragments fractionated by size using gel permeation HPLC (Fig. 1). Amino acid analyses of performic acid-oxidized aliquots showed that peaks A, B, C and D contained cysteic acid. SDS polyacrylamide gel analysis showed that peak A contained a mixture of partial cleavage products. Aliquots of material in the incompletely resolved peaks B and C were fully reduced with dithiothreitol and component polypeptides analysed by gel permeation HPLC

Table 1 Amino acid composition of disulphide-containing peptides isolated by HPLC from proteolytic digest of BHA cyanogen bromide fragments

	A	B	C	D	E	F	G	H
Cys	1.8 (2)	1.7 (2)	1.7 (2)	2.0 (2)	1.8 (2)	2.0 (2)	1.4 (2)	1.8 (2)
Asp	2.9 (3)	2.7 (2)	3.3 (3)	2.8 (2)	4.7 (4)	2.0 (2)	1.3 (1)	4.2 (4)
Threonine	—	4.6 (5)	0.2 (0)	4.6 (5)	0.8 (1)	—	1.1 (1)	—
Ser	1.6 (2)	1.3 (1)	0.8 (1)	0.6 (0)	1.0 (1)	—	—	2.6 (3)
Glu	1.8 (2)	3.5 (3)	1.0 (1)	2.6 (2)	2.9 (3)	1.2 (1)	2.0 (2)	2.5 (2)
Pro	1.2 (2)	4.1 (4)	1.0 (1)	4.3 (4)	—	—	—	2.2 (2)
Gly	2.6 (3)	2.5 (2)	3.0 (3)	3.1 (2)	3.2 (3)	—	1.4 (1)	2.8 (3)
Ala	0.9 (1)	1.2 (1)	1.0 (1)	1.3 (1)	1.2 (1)	—	—	1.8 (2)
Val	0.9 (1)	2.2 (2)	—	1.2 (1)	—	—	—	0.7 (1)
Met	—	—	—	0.1 (0)	—	† (1)	—	—
Ile	0.7 (1)	2.1 (2)	0.9 (1)	2.5 (2)	1.9 (2)	—	—	—
Leu	4.5 (5)	3.8 (3)	3.2 (3)	4.4 (3)	1.6 (1)	—	1.6 (2)	2.1 (2)
Tyr	—	0.6 (1)	—	—	1.2 (2)	—	—	1.6 (2)
Phe	—	1.0 (1)	—	1.1 (1)	1.7 (2)	—	—	2.9 (3)
His	—	2.0 (2)	—	2.0 (2)	1.9 (2)	—	—	—
Lys	—	3.2 (3)	—	2.0 (2)	1.1 (1)	—	1.0 (1)	1.0 (1)
Arg	1.0 (1)	—	0.9 (1)	0.1 (0)	0.3 (0)	—	—	1.0 (1)

Numbers in parentheses indicate predicted composition.
† Not quantifiable, present as oxidized homoserine lactone and homoserine.

0028–0836/81/050422—03$01.00

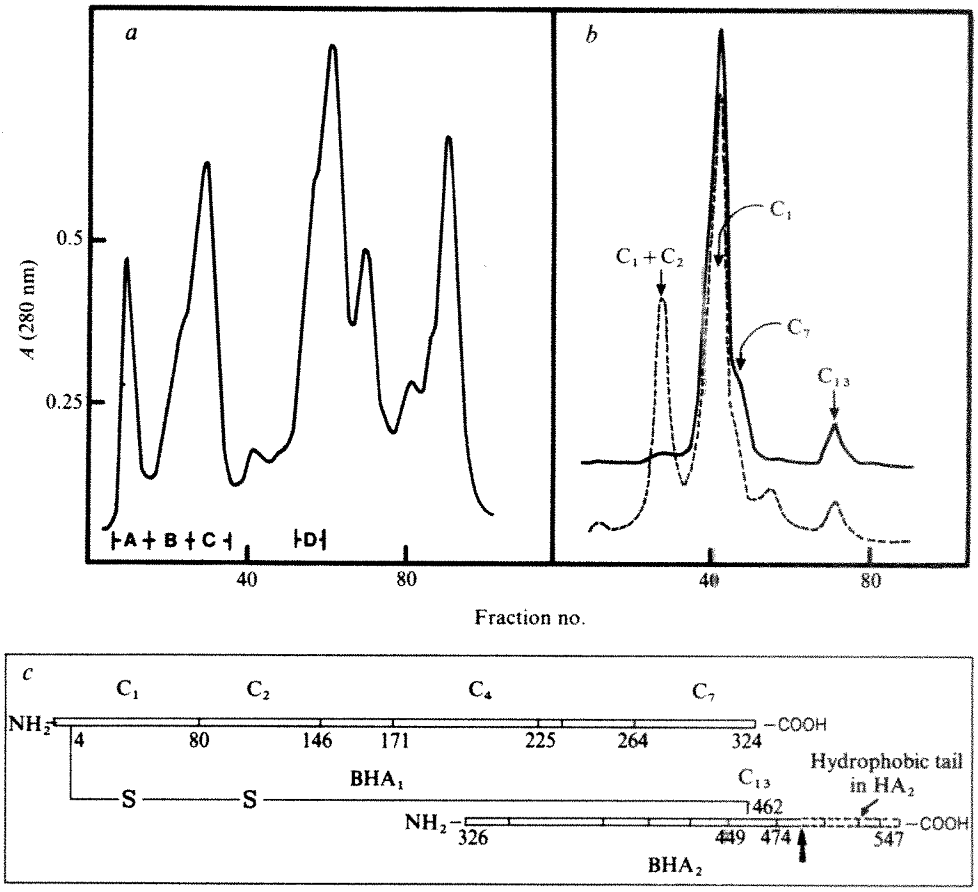


Fig. 1 Fractionation of the CNBr fragments of BHA by gel permeation HPLC. *a*, Separation of peptides from intact BHA. *b*, Separation of peptides released from peak B (——) and peak C (– – –) after full reduction and alkylation. BHA (100 nmol) was prepared from virus grown in the allantoic cavity of eggs[12], cleaved with CNBr in 70% formic acid for 24 h and directly injected in 4 aliquots onto two Waters I–125 columns run in series using a Waters M6000A pump and M450 detector (Milford). Columns were run in 6 M urea containing 0.2 M formic acid at 0.5 ml min^{-1} and monitored for absorbance at 280 nm; 0.5-ml fractions were collected and combined to give peaks A, B, C and D (*a*). Material in peaks B and C was deslated by dialysis or by absorption to Sep-Pak cartridges[13], fully reduced and alkylated in 6 M guanidine at *p*H 8.5 (ref. 12) and directly re-injected onto the gel permeation columns (*b*). Fractions corresponding to absorbance peaks were pooled, urea removed, and amino acid compositions and amino-terminal sequences determined using methods described elsewhere[7–9,12]. CNBr peptides C_1, C_2, C_7 and C_{13} were identified. *c*, Line drawing of the location of sites of CNBr cleavage of HA and the numbering of the half-cystines in fragments referred to in the text. The arrow indicates the approximate site of cleavage by bromelain. For the sake of clarity the original nomenclature[7] has been changed. Thus, here the partial cleavage product (C_1+C_2), and peptides C_1, C_2, C_4, C_7 and C_{13} correspond to BHA_1-C_1, BHA_1-C_2, BHA_1-C_4, BHA_1-C_3 and BHA_2-C_4, respectively, in Waterfield *et al.*[7].

(Fig. 1*b*). Appropriate fractions were pooled and the amino acid composition and amino-terminal sequence determined. Previous studies on the cyanogen bromide (CNBr) cleavage products of BHA_1 (refs 7, 8) and BHA_2 (ref. 9), knowledge of the complete amino acid sequence[3] and amino-terminal sequence analysis using methods described elsewhere[7,9] allowed us to establish the identity of the disulphide-linked fragments. Figure 1*c* shows the location of CNBr cleavages. Previous studies have shown that a partial CNBr cleavage occurs between residues 80 and 81 at a Met–Glu bond. This partial cleavage results in formation of a fragment containing C_1 joined to C_2. Peak B contained the partial cleavage fragment (C_1+C_2), fragment C_1 (which results from complete cleavage), the C-terminal fragment C_7 of BHA_1 and fragment C_{13} of BHA_2. Complete cleavage had occurred in the material eluted in peak C. After full reduction and re-chromatography, peak C was shown to contain peptides C_1 and C_7 of BHA_1 and C_{13} of BHA_2 (Fig. 1*b*). Peak D was found to contain a mixture of fragments C_2 and C_4 of BHA_1. Fragment C_4 is known to contain no half-cystine residues.

These results suggested that peptide C_2 of BHA_1 contained an intrachain disulphide bond and C_1 and C_7 of BHA_1 were linked to C_{13} in BHA_2.

CNBr fragment B was digested with *Staphylococcus aureus* V8 protease and the whole digest fractionated by reverse-phase HPLC on a C8 column (Fig. 2*a*). Fractions corresponding to optical density peaks were pooled, aliquots oxidized with performic acid and amino acid analyses carried out to locate half-cystine-containing peptides. A second portion of the V8 digest was further digested with trypsin and chromatographed and analysed in identical conditions (Fig. 2*b*). The amino acid compositions of pooled fractions containing half-cystine are given in Table 1 and the peptides can be located in Fig. 3 by the residue numbers of the half-cystine residues.

Two V8 peptides (A and B) which eluted at the end of the reverse-phase column run were identified by amino acid composition, using the single letter code of Dayhoff[10], as

(A) LGDCSIAGWLLGNPECDRLLSVPE
55 67

301
CPKYVKSE
|
(B) TKCQTPLGAINTTLPFHNVHPLTIGE
277

Further digestion with trypsin allowed identification of a tryptic cleavage fragment (C) which was derived from the V8 peptide (A). The V8 peptide (B) was recovered as peptide (D) after tryptic cleavage:

(C) LGDCSIAGWLLGNPECDR
55 67

301
CPK
|
(D) TKCQTPLGAINTTLPFHNVHPLTIGE
277

It was also possible to identify the following peptide (E)

4
(E) DQICIGYHANNSTE
|
LGNGCFEFYHK
462

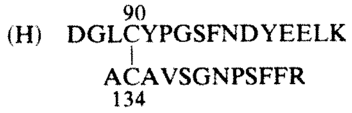
Re-chromatography of the digest using a flow rate of 0.5 ml min^{-1} allowed resolution of the peptides eluted early and made it possible to identify (F) and (G):

469
(F) CDDE
|
CM
473

42
(G) LCK
|
GTLENCE
273

These peptides allowed assignment of the single interchain bond which joins half-cystines 4 of BHA_1 and 462 of BHA_2 and also the assignment of intrachain disulphides between half-cystine residues 42 and 273, 55 and 67, 277 and 301 of BHA_1 and between 469 and 473 in BHA_2.

Assignment of the intrachain disulphide bond in CNBr peptide C_2 of BHA_1 was made by fractionation of a tryptic digest of a mixture of C_2 and C_4 which had been purified as peak D by gel permeation HPLC. The elution profile is shown in Fig. 2*c* and the composition of peptide (H), recovered in peak H, is given in Table 1. This compositional analysis confirmed the location of a disulphide bond between half-cystines 90 and 134:



$$\text{(H)}\ \begin{array}{l}\text{DGL}\overset{90}{\text{C}}\text{YPGSFNDYEELK}\\ \quad\ \ |\\ \text{A}\underset{134}{\text{C}}\text{AVSGNPSFFR}\end{array}$$

The peptides isolated from BHA allow the assignment of disulphide bonds which link the twelve half-cystine residues of BHA. The results are summarized in Fig. 3. A similar study of the disulphide bonds of the haemagglutinin from Hong Kong virus A/Memphis/102/72 (H3N2) has been reported by Dopheide and Ward[11] who present evidence that half-cystine 139 of HA_1 is linked to half-cystine 472 in HA_2 in the H3 strain and suggest that residue 97 may be linked to residue 476. Our results show that half-cystine residues homologous to those at positions 139, 97, 472 and 476 are arranged differently in the H2 virus HA. Because the position of half-cystine residues is highly conserved in haemagglutinins of viruses of both subtypes[3] it is unlikely that the disulphide bonds differ. Analyses of the three-dimensional structure of the BHA fragment from an H3 virus (X-31) indicates a single interchain bond[5] and it is likely that methodological difficulties remain to be resolved in the analysis of the disulphide-linked thermolytic peptides isolated by Dopheide and Ward[11]. In agreement with their results, we assign an interchain bond between half-cystine residues 14 of HA_1 and 465 in HA_2 in the H3 strain, which are homologous with residues 4 of BHA_1 and 462 in BHA_2 in the H2 strain. Our results also agree with those of Dopheide and Ward[11] in the placement of intrachain disulphides between half-cystines at residues 52 and 277, 64 and 76, and 281 and 305 in HA_1 of the H3 strain, which are homologous to disulphides in HA_1 of the H2 strain which link residues 42 and 273, 55 and 67 and 277 and 301.

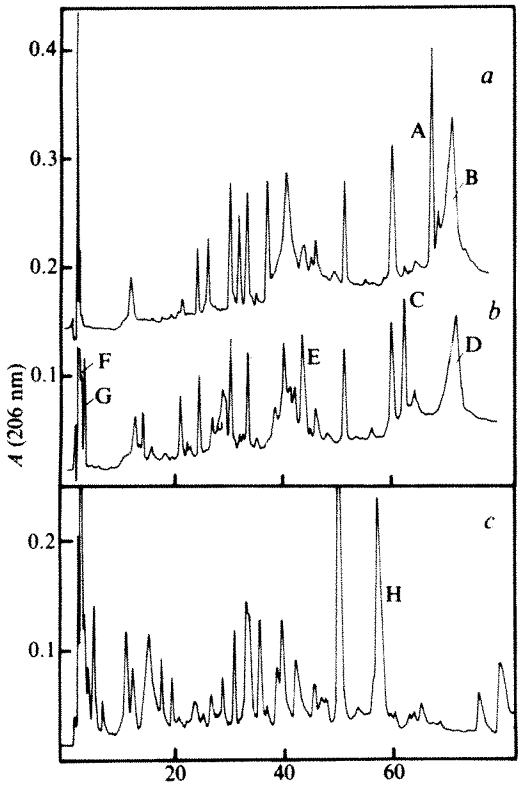
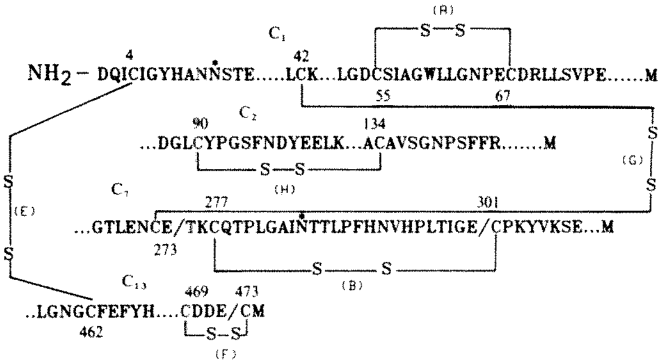


Fig. 2 Separation of enzymatic digests of disulphide-containing CNBr fragments by reverse-phase HPLC. Fragment C (purified as described in Fig. 1 legend) was digested with *S. aureus* V8 protease (70 nmol, enzyme/substrate 1:18) in 0.1 M ammonium bicarbonate buffer for 18 h and directly injected onto a Dupont Zorbax C8 column which was eluted with a gradient of acetonitrile in 10 mM ammonium acetate buffer, *p*H 6.5 (ref. 13). *a*, Column effluent was monitored for absorbance at 280 and 206 nm using two Uvicord S monitors (LKB Instruments) equipped with HPLC flow cells. Fractions (0.75 ml) were collected, peaks pooled by optical density and aliquots oxidized with performic acid and the amino acid composition determined. *b*, An aliquot of the V8 digest was further digested with trypsin (enzyme substrate 2% by weight) and analysed by the same methods. Material from pool D (see Fig. 1) was digested with 2% trypsin (by weight) and analysed by chromatography on Zarbox C8 as described for pool C. *c*, Separation of tryptic peptides. Peptides with disulphide bonds are indicated by letters referred to in the text.



(R)
S—S
C1
4
42
NH2 – DQICIGYHANNSTE.....LCK...LGDCSIAGWLLGNPECDRLLSVPE......M
55
67
C2
90
134
...DGLCYPGSFNDYEELK...ACAVSGNPSFFR.......M
S—S
(H)
S
(G)
S
S
(E)
S
C7
277
301
...GTLENCE/TKCQTPLGAINTTLPFHNVHPLTIGE/CPKYVKSE...M
273
S — S
(B)
C13
469
473
..LGNGCFEFYH....CDDE/CM
462
S–S
(F)

Fig. 3 The disulphide bonds of the bromelain-released fragments of the haemagglutinin of A/Japan 305/57 (H2N1). * Indicates the attachment site for carbohydrate side chains.

The results presented here establish disulphide bonds for the region of HA which is external to the membrane of the virus. A further four half-cystine residues are located near the C terminus of the HA of A/Japan 305/57(H2N1)—two of these are conserved in the HAs of all strains examined (543 and 546)[3]. These half-cystines may be located inside the virus particle, below the lipid bilayer because N-terminal to these residues there is a 26–30-residue hydrophobic sequence that would be expected to cross the lipid bilayer. The remaining two half-cystines of this HA, which are not conserved in other strains, may simply serve as hydrophobic or uncharged amino acides. Whether these four half-cystines are disulphide bonded remains to be established. The contribution of the disulphide bonds to the three-dimensional structure of the HA is described by Wilson *et al.*[6].

Received 30 October; accepted 23 December 1980.

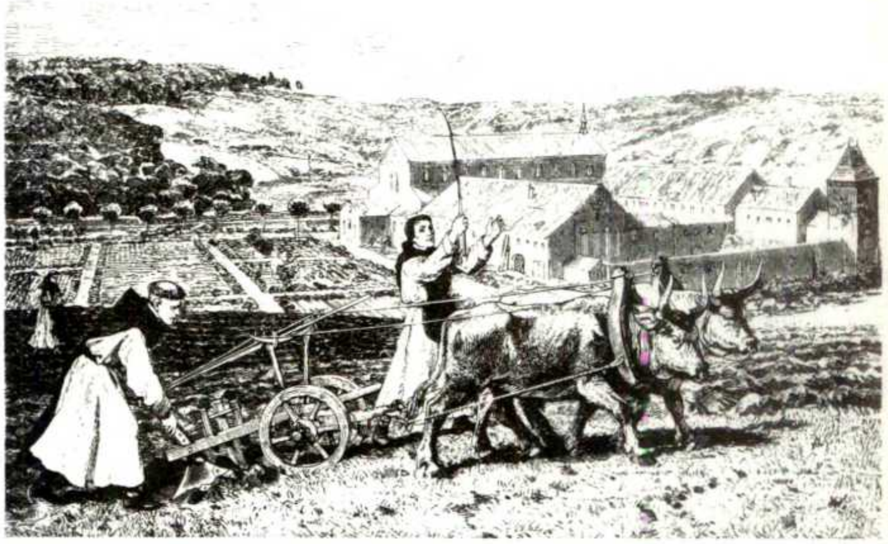
1. Porter, A. G. *et al. Nature* **282,** 471–477 (1979).
2. MinJou, W. *et al. Cell* **19,** 683–696 (1980).
3. Gething, M. J., Bye, J., Skehel, J. J. & Waterfield, M. D. *Nature* **287,** 301–306 (1980).
4. Brand, C. M. & Skehel, J. J. *Nature new Biol.* **238,** 145–147 (1972).
5. Wilson, I. A., Skehel, J. J. & Wiley, D. C. in *Structures and Variations in Influenza Virus* (*Developments in Cell Biology 5*) (eds Laver, W. G. & Air, G.) 339–384 (Elsevier, Amsterdam, 1980).
6. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289,** 366–373 (1981).
7. Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. *Br. med. Bull.* **35,** 57–64 (1979).
8. Waterfield, M. D., Skehel, J. J., Nakashima, Y., Gurnett, A. & Bilham, T. in *The Influenza Virus Haemagglutinin* (*Topics in Infectious Diseases*) Vol. 3 (eds Laver, W. G., Bachmayer, H. & Weil, R.) 167–169 (Springer, Vienna, 1978).
9. McCauley, J., Skehel, J. J. & Waterfield, M. D. in *The Influenza Virus Haemagglutinin* (*Topics in Infectious Diseases*) Vol. 3 (eds Laver, W. G., Bachmayer, H. & Weil, R.) 81–192 (Springer, Vienna, 1978).
10. Dayhoff, M. O. (ed). *Atlas of Protein Sequence and Structure* Vol. 5, D2 (National Biomedical Research Foundation, Maryland, 1972).
11. Dopheide, T. A. A. & Ward, C. A. *FEBS Lett.* **110,** 181–182 (1980).
12. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 93–97 (1975).
13. Waterfield, M. D. & Scrace, G. T. in *Biological/Biomedical Applications of Liquid Chromatography III* (ed. Hawk, G.) (in the press).

BOOK REVIEWS

Swimming against the tide

Eric Ashby

The Politics of Self-Sufficiency. By Michael Allaby and Peter Bunyard. Pp.242. (Oxford University Press: 1980.) Hbk £7.95, $24.95; pbk £3.95.



Self-sufficiency in the Middle Ages — how viable today?

IF YOU see someone swimming against the tide, watch him. He may sink without trace; he may bring ashore a new idea. It is in this spirit that I commend this duet by Michael Allaby and Peter Bunyard. A duet? Yes, for in each chapter first one author, then the other, has his say. The two authors agree in principle that self-sufficiency in a Good Thing. They do not agree about the way in which self-sufficiency might be achieved nor about the part it might play in Western industrial society.

On the first page the authors define what they are writing about. It is not the self-sufficiency of pioneers colonizing a hostile environment, often by a ruthless exploitation of nature; it is self-sufficiency as an alternative "lifestyle": the community which grows its own food (preferably without artificial manures), generates its own energy (or some of it) from sunlight and the burning of timber, disposes of its own wastes, imports such things as it cannot make in exchange for its own surplus, by barter, not money. There are already little communities of this kind in Britain, and the Israeli kibbutz and Chinese commune are viable versions of a similar idea. Indeed, if one is looking for historical precedents for self-sufficiency, the mediaeval monastery is surely an impressive example.

The self-sufficient community, like the monastic community, can be sustained as a minority way of life in an industrial pluralistic democracy. It is when enthusiasts for self-sufficiency urge us to make it a majority way of life that doubts arise. All sorts of awkward questions have to be answered. Is it, in any case, a "better" lifestyle? Could such a lifestyle be sustained by the minorities who enjoy it if they were not embedded in a conventional capitalist society which can come to their rescue when they need to travel, to listen to music, to get medical attention, to type and publish books such as this one, to replace tools? And, since this is a book about the politics of self-sufficiency, under what kind of political system could self-sufficiency ever become a majority lifestyle?

The virtue of the book is that it does examine these questions honestly. It considers the prospects for a mass-movement toward self-sufficiency. The prophets have to step warily between the ideologies of anarchy on one side and fascism on the other. Except for hermits, self-sufficiency implies co-operation, and "the moment a group of people act collectively they cease to act anarchically". So the self-sufficient society would have to submit itself to some sort of government. If it is to be a government which rejects an economic system based on motives of profit, then "the economy itself would be changed from a market to a command type. The injunction to industrialists would no longer be the simple 'Be profitable!' but the much more complex 'Make this, but not that'." The only known patterns of government capable of behaving this way are far-left or far-right dictatorships. The authors don't relish either of these. They seek a middle way: some communal ownership of property; some concession to the fact (which they freely admit) that an atavistic retreat to a bucolic past is simply not on; some compromise with the harsh reality of the human condition, which is that greed, selfishness and love of comfort are majority attitudes, and that beneficence, self-denial and austerity are likely to remain minority attitudes. They adduce plenty of evidence that small communities can — and some do — drastically reduce their reliance upon the products of industrial society, and consume much less per capita than the average consumption of the limited resources of nature. They make the point clearly that they do not consider self-sufficiency to be a middle class élitist indulgence or a highbrow name for drop-outs. Self-sufficiency is a deep — they call it a "spiritual" — protest against a way of life which is carrying us out on a tidal current to God knows what dangers.

Anyone who seriously contemplates self-sufficiency in political terms had better stand at Waterloo Station in the rush hour and ask himself exactly how any political movement is going to lead this surge of commuters into such a different lifestyle, and ensure that they are adequately fed, sheltered and kept warm in (say) the idyllic conditions of rural Devon and Cornwall. (And, incidentally, ask himself what would happen to Devon and Cornwall if such a miracle occurred.) But this would be a slick and insensitive way to end this review: for there is one important conclusion to be drawn from this book, though it may not be a conclusion the authors approve. It is that little self-sufficient communities such as Allaby and Bunyard describe have an important part to play in contemporary Britain. Like monasteries in the Dark Ages, they affirm that there is another, and perhaps a better, way to live than the way most people live and that it is not out of reach. They are not composed of cranks. They are men and women who believe in swimming against the tide. They may sink without trace. But, like monasteries in the Dark Ages, they may carry the seeds for a new kind of civilization, one able to transform the taint of unemployment into the boon of leisure for those without work, and the drudgery of attending an assembly line for the satisfaction of manual labour for those who have work. □

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Health risk: just a guessing game?

T. A. Connors

Ethylene Dichloride: A Potential Health Risk? 5th Banbury Report. Edited by B. Ames, P. Infante and R. Reitz. Pp.350. (Cold Spring Harbor Laboratory: 1980.) $45, $54 outside USA.

"WELL, *is* ethylene dichloride a health risk?" is the question that comes to mind after reading this book. The comment does not imply any specific criticism of the report but rather sums up our present knowledge of the toxicology of a wide variety of chemicals to which human beings are exposed. We suspect from animal and other laboratory tests that many chemicals may be a risk to human health as carcinogens, teratogens, mutagens and so on, but we do not have reliable methods to determine whether there is an actual risk in any situation or what the degree of that risk may be.

Ethylene dichloride is a very important chemical indeed, the US production in 1977 amounting to 11,000 million pounds in weight. It is used mainly as an intermediate, so most high-level exposure is in the chemical industry, but significant amounts are also used in the textile and food industries and in agriculture, for example. The possibility that ethylene dichloride might be carcinogenic was first suggested when it was shown to be mutagenic in bacteria. In a large international study, tests of this type have been shown — at least for certain classes of chemical — to correlate reasonably well with animal carcinogenicity tests which in turn have given positive results for most human carcinogens identified by epidemiological or case control studies. This Banbury Report summarizes what is known of the toxicology of ethylene dichloride and one learns that it can be metabolized *in vivo* to an alkylating agent which can react with DNA and initiate the events that many people think can lead to cancer. One also learns that other short-term tests, using *Drosophila* and yeast, for example, have given positive results and that the chemical, like many of its congeners, is carcinogenic in rats and mice. However not all of the data are unequivocal. In a well-controlled study, Maltoni and his colleagues could find no evidence for the carcinogenicity of ethylene dichloride in rats and they question the validity of earlier studies from a number of aspects including "the professionality of the team carrying on treatment, control of animals and autopsies". Furthermore, no teratogenicity or major reproductive toxicity was found either in animals or in exposed workers, nor was there any evidence of an increased incidence of cancer in workers exposed to ethylene dichloride for 20 years or more.

All in all the book is a wonderful miniature of the multidisciplinary field of toxicology and the problems that arise in interpreting data from different types of test, in choosing between conflicting results and in attempting to assess the risk to human beings. Although scientists, as in this report, are to be commended for attempting to measure carcinogenic potency and to quantify risk, the equation contains so many unknowns that the assessed level of risk must be seen to be based on a working hypothesis with many assumptions rather than as the result of an extrapolation based on scientifically reliable methods. It follows then that in making the decision to restrict a chemical, to replace it with "safer" alternatives or to ban it altogether, economic and social argument must be considered as well as the scientific aspects. There is concern that a number of chemicals which have been used for many years and considered to be safe because there was no acute toxicity associated with exposure, may nevertheless be highly dangerous in the long term. The Banbury Report, with contributions from industrialists, government scientists and university researchers, shows how widespread this concern is, but until animal experiments can be related to human situations the debate on the dangers of chemicals will continue.

However, there has already been one welcome outcome. Workers are nowadays exposed to less dirty, dusty and smelly conditions than even a few years ago, while urban environments are certainly more wholesome than they once were. Whatever the effects may be on the cancer incidence only time will tell; but for those people who believe that quality of life is at least as important as quantity of life then there is already a vast improvement. □

T.A. Connors is at the MRC Toxicology Unit, Carshalton, Surrey.

Catastrophes for starters

Colin Upstill

An Introduction to Catastrophe Theory. By P.T. Saunders. Pp.144. (Cambridge University Press: 1980.) Hbk £9.50, $27.50; pbk £3.25, $8.95.

ALL previous introductions to catastrophe theory have been either popularizations, devoid of any serious mathematics and often riddled with falsehoods, or works of considerable mathematical sophistication. This book succeeds in filling the gap between these extremes. It divides naturally into two parts. The first is an exposition of the theory, which deals with the singularities of smooth real valued functions, using only a limited mathematical vocabulary yet not glossing over anything of importance: structural stability, equivalence, the splitting lemma, determinism and universal unfoldings are all explained as clearly as one could wish. The second half of the book gives the flavour of applications of the theory by way of a selection of examples from the physical, social and biological sciences, all of which are to be found discussed in much greater detail elsewhere in the literature. The author is careful not to exaggerate the status of catastrophe theoretic models in the social and biological sciences, and concludes with a lucid discussion of the explanatory powers of the theory and some cautionary words on appropriate standards of judgement of applications in different disciplines.

My enthusiasm is not unqualified, however. Central to catastrophe theory are the geometries of the catastrophes themselves, so it is inexcusable that their illustration in this book leaves something to be desired. Most of the relevant figures are slavishly and uncritically copied, with due acknowledgement, from the book by Bröcker and Lander (*Differentiable Germs and Catastrophes*; Cambridge University Press, 1975); some include labelling in a nomenclature incompatible with that used in Saunders's text, others include construction lines which are here irrelevant and unexplained, or fail to display the nature of the singularity at the origin of control space. Where the drawings are original, similar lack of care is evident. At the hyperbolic umbilic singularity, the rib (cusped edge) and the fold touch parabolically, yet here they are depicted as intersecting straight lines; the sketch of the lips event is quite dreadful.

The text does not suffer from any such sloppiness, but there are some omissions which are surprising in such a recent publication. Most notable is the absence of any reference to the work of the Russian mathematician V.I. Arnol'd, who is responsible for the enormous extension of Thom's classification to catastrophes of codimension >4, and for a rational system of symbols to label the catastrophes, which Saunders eschews — the Thomist "pet names" are fine if one is only considering a few low-dimensional catastrophes, but things get out of hand as the dimensionality and number of catastrophes in the list increases. Explanations of the role of higher dimensional catastrophes as organizing centres and of the importance of compact catastrophes in applications would not come amiss; neither would some discussion of (or at least references to) recent vociferous criticisms of some applications of catastrophe theory in the social and biological sciences.

In a rapidly expanding subject it is inevitable that any book will be out of date as soon as it is published. However, much of this book could have been written four years ago. Whilst there is no claim to define the "state of the art", it is an uneven treatment, to say the least, that for some applications covers material from one or two expository works of several years ago, but for others refers to a quantity of very specialized material of recent date. One is forced to conclude that the author has made no great effort to acquaint himself with recent work outside his own particular interests, which is a pity, because as a result he has written a good book when he could have written a superb one.

Colin Upstill is a Research Fellow at the H.H. Wills Physics Laboratory, University of Bristol.

Potpourri of nuclear structure theory

J. M. Irvine

Nuclear Spectroscopy. Lecture Notes in Physics, 119. Edited by G. F. Bertsch and D. Kurath. Pp.250. (Springer-Verlag: 1980.) DM33, $19.50.

THE Gull Lake workshop on nuclear spectroscopy, held in the summer of 1979, reviewed most of the basic topics in theoretical nuclear structure physics. This book is a record of the deliberations of the participants.

The opening chapter by Gerry Brown takes us through the familiar analysis of the meson exchange nature of nuclear forces, linking it to the quark structure of the nucleon in the "little bag model" of Brown and Rho. The Brueckner–Bethe and Landau–Migdal theories of effective interactions are developed and the predictions for the bulk properties of nuclear matter presented. In other interesting papers, Dieter Kurath presents an introduction to the nuclear shell-model which is novel in that it includes a short discussion of the current shell-model computer codes, and George Bertsch discusses nuclear vibrations — both in the random phase approximation and Landau formalisms — and presents an up-to-date account of the new giant resonances. Amand Faessler gives a unified account of deformed and transitional nuclei. Concentrating first on high spin states in deformed nuclei, he discusses the "backbending" phenomena and the nature of Yrast traps, and goes on to give an account of the breakdown of pairing due to the Coriolis force in transitional nuclei. Franco Iachello describes the use of group theory in the analysis of nuclear spectroscopy with a natural emphasis on the successes of the interacting boson approximation.

The use of statistical analysis in describing complex spectroscopies has been advocated for a long time. At last there are signs that it is being put to use in nuclear physics and it is appropriate that the field should here be reviewed by J. B. French. Finally, there is an interesting appendix by Bertsch, Zamick and Mekjian which highlights some of the unsolved theoretical puzzles which abound in nuclear structure physics.

I found this an exceptionally readable account of current topics in nuclear theory. The level of presentation of the material is remarkably uniform for a multi-author volume. I am sure that students and research workers in the field will find this a useful addition to their bookshelves. □

J. M. Irvine is Reader in Theoretical Physics at The University of Manchester.

Various vectors and plant diseases

G. D. Heathcote

Vectors of Plant Pathogens. Edited by Kerry F. Harris and Karl Maramorosch. Pp.467. (Academic: 1980.) £27, $48.

THIS book succeeds admirably in what it sets out to do, providing much detailed, up-to-date information on the interrelationships of many plant pathogens, their plant hosts and various vectors. It is third in a series of four volumes being published on this general topic. The earlier volumes were concerned with aphids and leaf-hoppers, and, although the first chapter of this volume is devoted to pathogen–vector relationships within these groups of insects, most of the current book is concerned with other, generally less well-known vectors. These range from mealybugs, to fungi and to nematodes where, although the relationships may be equally close and complex, they can be non-obligate and accidental, as with dipterous flies and many bacterial plant pathogens.

The book clearly points out the wide gaps in our knowledge and understanding of this subject. Many chapters end with questions for the reader or indicate need for further study. One contribution is even devoted to the search for the vector of a disease (lethal yellowing of coconut palm) where the very nature of the pathogen is still uncertain, although the reader may assume it to be a mycoplasma-like organism (MLO); a cixiid bug appears to be the most likely candidate as the vector.

The 13 chapters are written by different specialists, or groups of specialists who were clearly given a free hand and who have very different styles of writing. This variation in style adds to the interest but perhaps the editors might have tied the volume together a little more closely.

This thought-provoking book contains many references and will be of value and interest to researchers as well as to students and teachers. For example, how many of us have considered the possibility of a link between the use of mustard gas in the First World War and the recent spread of Dutch elm disease? The book suggests one. □

G. D. Heathcote is at Broom's Barn Experimental Station, Higham, Bury St Edmunds, Suffolk, and works on transmission of viruses by aphids.

Seabird synthesis

J.B. Nelson

Behavior of Marine Animals. Vol.4, Marine Birds. Edited by J. Burger, B.L. Olla and H.E. Winn. Pp.515. (Plenum: 1980.) £28.35, $54.

STUDENTS of seabirds will warmly welcome this fine and wide-ranging compilation of data, ideas and references. Many important aspects of breeding biology such as choice of habitat, mate-selection, various factors and strategies relating to breeding success and the nature of communication behaviour are excellently reviewed. These articles form the core of the book but three others (seabirds at sea, the distribution of gulls in North America and, somewhat oddly placed, chemoreception in seabirds) stand outside this major theme.

Systematic use of headings and subheadings produces welcome internal consistency within the volume but, in view of inevitable omissions in such a wide-ranging and essentially review work (references average 148 per article, excluding that by Southern), space could have been saved by editorial restrictions on the text authors devote to their own research. As it is, several contributors suddenly abandon the review style and launch into detailed and specialized exposition. Undoubtedly, though, seabird workers will warmly welcome this work as evidence that their research, despite its heterogeneous and incomplete nature, nevertheless engenders valuable syntheses. Seabird studies, as one might expect after two highly productive decades, are moving away from general data gathering and towards more detailed investigations of specific problems. These contributions point the way to many such areas. Perhaps the chief merit of the book lies in its discussion of the relevance of this mass of work to well-recognized problems, and in the clarification of these problems, rather than in the formulation of new concepts. Old terms, hitherto loosely and confusingly employed — for instance synchrony, colony, fledging, starvation period, display, social stimulation and the Darling effect — are carefully dissected. Again, nowhere else in all the seabird literature can one find a properly detailed discussion of the effect of age on breeding success. As an example of "properly detailed" there is Burger's Table II, which gives information individually for 86 species under 10 subject headings.

Yet there are fundamental concepts, affecting most of these subjects, which receive no critical attention. I would have welcomed a thorough review of the concept of direct, density-dependent competition which I am convinced often does not operate among seabirds. Within the constraints of this volume, no single contributor could have tackled the subject of the regulation of seabird populations, but there are reliable and reviewable field data which bear directly on the validity of density-dependent regulation in seabirds.

Among the many ideas canvassed, I liked Hunt's emphasis on the need to distinguish a lifetime's reproductive output from that of any one year, an important but rather inaccessible statistic when considering the reproductive strategies so beloved of today's theoreticians. Ryder's excellent summary of age-effects brings out the intriguing phenomenon of reduced success in old individuals. In related vein, Montevecchi and Porter indicate the minimum detail which must be obtained on time and energy budgets before one can begin to theorize about optimal use of resources. They emphasize, as does Gochfield, what I too have suggested, that for many seabirds, especially males, the costliest part of the reproductive cycle is not chick-rearing but the pre-laying stage when sites are established (or re-established) and courtship performed. Beer, bristling with ideas, emphasizes the importance of context on the form and function of display — a minefield of potential ambiguity. His article should be required reading for behaviourists, to steer them away from beguilingly simplistic approaches. But even he founders in the complexities inherent in the notion of "intention" in animals.

Gochfield must speak for many in his plea for adequate presentation of data (and he gives an excellent account of methods) to allow others to re-work them. Highly processed results may satisfy statisticians but are useless, except in a take-it or leave-it form, to the Indians.

Bridge-building is an important function of review articles and Brown diligently and perceptively connects the many scattered observations on birds at sea with some oceanographic data. Even so, the key words remain "broad outline" in his statement that "we now know in broad outline the pelagic distribution of the common seabirds in summer at least, in almost all the world's oceans". Many of his valuable references will be new to most ornithologists.

Throughout the book, and rightly so, functional interpretations are stressed, for it is essential to generate testable ideas about the adaptive significance of the intricate web of behaviour by means of which animals cope with environmental demands. ☐

J.B. Nelson is Reader in Zoology at the University of Aberdeen, and the author of Seabirds: Their Biology and Ecology, *published earlier this year by Hamlyn.*

The second of the six core volumes of the *Handbook of Applicable Mathematics* has been published by John Wiley (for review see *Nature* **288**, 37). Volume II, *Probability*, edited by Emlyn Lloyd, is available as a single volume at £32.50, $85. Subscription price for the core volumes is £27.50, $75.

Mass spectrometry for biochemists

Howard R. Morris

Biochemical Applications of Mass Spectrometry: First Supplementary Volume. Edited by George R. Waller and Otis C. Dermer. Pp.1,279. (Wiley-Interscience: 1980.) £79.80, $187.50.

THIS is the first supplementary volume of a previous work of the same title, and is a compendium of contributions by well-known authors in specialist fields. The chapters have been written at a level suitable for postgraduate students, and in general they provide useful background reading for the "layman", whether he be a newcomer to research in one of these fields, or a biochemist or biologist seeking a solution to a particular problem.

The editors predict in the preface that certain sections are unlikely to be comprehensive in content, and this is indeed so in a number of cases. For example, in the section on instrumentation there is, inexcusably, no reference to one of the most significant and highly relevant developments in the area of biochemical mass spectrometry, the high-field magnet. The new range of commercial instruments based upon this concept (giving high sensitivity operation up to m/e 1800 — "extended" — or m/e 3000 — "high-field") is of special interest and importance to the biochemist, since he is often dealing with small amounts of high-molecular-weight polar compounds.

A considerable portion of this large book is taken up with matters which are not *per se* biochemical mass spectrometry, such as basic principles, and a listing of the equipment in several different spectrometry laboratories. The choice of these laboratories would appear to be somewhat arbitrary, since a number of them are not noted for major contributions in the field of biochemistry. The validity of detailed descriptions of mass spectrometric data systems is in any event debatable for the intended audience. Similarly, in the section on interpretation of mass spectra, a series of methods and computerized techniques is given which, as far as one can tell, have not yet been applied to the solution of biochemical problems, and would therefore more properly have appeared in a subsequent volume.

In contrast, the section covering applications makes very worthwhile reading, since this is the real meat of biochemical mass spectrometry. The chapter on antibiotics is excellent, and that on proteins is commendably comprehensive, despite a few inevitable inaccuracies in the interpretation of some of the details of the work covered. Unfortunately, not all of the authors are so thorough; for example the section on tetrapyrroles is notable for its superficial coverage. Finally, the descriptions of some interesting new applications make enjoyable reading, in particular the extension of mass spectrometric techniques into the study of volatiles emitted by human beings; one might predict a rapid growth of interest in this area since the information provided would appear to be highly relevant to the physiological status of the body in both health and disease.

As a whole, *Biochemical Applications of Mass Spectrometry* is a useful and often stimulating reference work, providing the biochemical student or researcher with a basic background from which to make a more detailed exploration of the literature. ☐

Howard R. Morris is Professor of Biological Chemistry in the Department of Biochemistry at Imperial College, University of London.

BOOKS RECEIVED

Mathematics

ANDERSON, S. *et al.* Statistical Methods for Comparative Studies. Techniques for Bias Reduction. Pp.289. ISBN 0-471-04838-0. (Wiley: 1980.) £13.50.

GAL, S. Search Games: Mathematics in Science and Engineering, Vol. 149. Pp.216. ISBN 0-12-273850-0. (Academic: 1980.) $20.

STEWART, W. E., RAY, W. H. and CONLEY, C. C. (eds). Dynamics and Modelling of Reacting Systems. Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, The University of Wisconsin-Madison, October 1979. Pp.413. ISBN 0-12-669550-4. (Academic: 1980.) $27.50.

Astronomy

BALIAN, R., AUDOZE, J. and SCHRAMM, D. N. (eds). Physical Cosmology. Proceedings of Les Houches Summer School, Session XXXII, July 1979. Pp.668. ISBN 0-444-85433-9. (North-Holland: 1980.) $109.75, Dfl. 191.

HALLIDAY, I. And McINTOSH, B. A. (eds). Solid Particles in the Solar System. Symposium No. 90 Organized by the IAU in co-operation with COSPAR held at Ottawa, August 1979. Pp.432. Hbk ISBN 90-277-1164-X; pbk ISBN 90-277-1165-8. (Reidel: 1980.) Hbk Dfl. 95, $49.95; pbk Dfl. 50, $26.50.

Physics

BERTOZZI, W., COSTA, S. and SCHAERF, C. (eds). Electron and Pion Interactions with Nuclei at Intermediate Energies. Studies in High Energy Physics, Vol.2. Proceedings of the International School sponsored by the University of Rome, held at the Centre Studi in Ariccia, Rome, June 1979. Pp.703. ISBN 3-7186-0015-3. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) $55, Dfl.108.

BERTRAND, F. E. (ed.). Giant Multipole Resonances. Nuclear Science Research Conference Series, Vol.1. Proceedings of the Conference held at the American Museum of Science and Energy, Oak Ridge, Tennessee, October 1979. Pp.482. Flexi ISBN 3-7186-0029-5. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) $42.50, Dfl.83.

GIBSON, W. M. and POLLARD, B. R. Symmetry Principles in Elementary Particle Physics. Pp.389. Hbk ISBN 0-521-20787-8; pbk ISBN 0-521-29964-0. (Cambridge University Press: 1980.) Hbk np; pbk £9.95.

HARGRAVES, R. B. (ed.). Physics of Magnetic Processes. Pp.585. Hbk ISBN 0-691-08259-6; pbk ISBN 0-691-08261-8. (Princeton University Press: 1980.) Hbk £21.90; pbk £8.40.

HOOFT, G. 't *et al.* Recent Developments in Gauge Theories. NATO Advanced Study Series. Series B: Physics Vol.59. Pp.438. ISBN 0-306-40479-6. (Plenum: 1980.) $49.50.

HOROWITZ, P. and HILL, W. The Art of Electronics. Pp.716. Hbk ISBN 0-521-23151-5; pbk ISBN 0-521-29837-7. (Cambridge University Press: 1980.) Hbk £35; pbk £12.50.

JERRARD, H. G. Electro-Optics/Laser International '80 UK Conference Proceedings, Brighton, March 1980. Pp.401. ISBN 0-86103-035-4. (IPC Science and Technology Press Ltd, Guildford, 1980.) £22.50.

KOLLMANN, W. Computational Fluid Dynamics, Vol.2. Pp.266. ISBN 0-89116-192-9. (Hemisphere Publishing: 1980.) $55.

PEEBLES, P. J. E. The Large-Scale Structures of the Universe. Pp.421. Hbk ISBN 08240-5. (Princeton University Press: 1980.) Hbk £16.50; pbk £5.50.

SCHUTZ, B. F. Geometrical Methods of Mathematical Physics. Pp. 250. Hbk ISBN 0-521-23271-6; pbk ISBN 0-521-29887-3. (Cambridge University Press: 1980.) Hbk £18; pbk £7.50.

TAMURA, T., NATOWITZ, J. B. and YOUNGBLOOD, D. H. (eds). Continuum Spectra of Heavy Ion Reactions. Nuclear Science Research Conference Series, Vol.2. Proceedings of the International Symposium, San Antonio, Texas, December 1979. Pp.476. Flexi ISBN 3-7186-0028-5. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) $48.50, Dfl. 96.

Chemistry

CARR, P. W. and BOWERS, L. D. Immobilized Enzymes in Analytical and Clinical Chemistry. Fundamentals and Applications. Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications, Vol.56. Pp.460. ISBN 0-471-04919-0. (Wiley: 1980.) £24.

CHAMOT, E. M. and MASON, C. W. Handbook of Chemical Microscopy, Vol. II. Chemical Methods and Inorganic Qualitative Analysis. 2nd Edn. Pp.438. ISBN 0-471-04122-X. (Wiley: 1980.) £14.75.

DE MAYO, P. (ed.). Rearrangements in Ground and Excited States, Vol. 3: Organic Chemistry, Vol. 42. Pp.607. ISBN 0-12481303-8. (Academic: 1980.) Np.

HEAL, H. G. The Inorganic Heterocyclic Chemistry of Sulphur, Nitrogen and Phosphorus. Pp.271. ISBN 0-12-335680-6. (Academic: 1980.) £35.80, $82.50.

HUTZINGER, O. (ed.). The Handbook of Environmental Chemistry, Vol.1. Part A. The Natural Environment and the Biogeochemical Cycles. Pp.258. ISBN 3-540-09688-4. (Springer: 1980.) DM 98, $57.70.

JOHNSON, B. F. G. (ed.) Transition Metal Clusters. Pp.681. ISBN 0-471-27817-3. (Wiley: 1980.) Np.

KRYT, D. (ed.). Dictionary of Chemical Terminology. In five languages: English, German, French, Polish and Russian. Pp.612. ISBN 0-4444-99788-1. (Elsevier Scientific: 1980.) $95, Dfl.195.

WEBB, G. A. (Sen. Reporter) Nuclear Magnetic Resonance. Vol.9: A Review of the Literature published between June 1978 and May 1979. Pp.336. ISBN 0-85186-960-2. (The Royal Society of Chemistry: London, 1980.) £48, $129.

Technology

BARD, A. J. and FAULKNER, L. R. Electrochemical Methods. Fundamentals and Applications. Pp.718. ISBN 0-471-05542-5. (Wiley: 1980.) £14.70.

BASELT, R. C. Biological Monitoring Methods for Industrial Chemicals. Pp.301. ISBN 0-931890-04-7. (Wiley: 1980.) £24. USA: Biomedical Publications, PO Box 368, Canton, Connecticut 06019.

BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE. Energy in the Balance. Some Papers from BA79 given at the Annual Meeting of the BAAS Heriot-Watt University, Edinburgh, 1979. Pp.234. Flexi ISBN 0-86103-031-1. (Westbury House: PO Box 63, Bury Street, Guildford, Surrey GU2 5BH, 1980.) Np.

DUFFIE, J. A. and BECKMAN, W. A. Solar Engineering of Thermal Processes. Pp.762. ISBN 0-471-05066-0. (Wiley: 1980.) £13.80.

GROSS, W. A. *et al.* Fluid Film Lubrication. Pp.774. ISBN 0-471-08357-7. (Wiley: 1980.) £18.65.

HUGHES, D. A Literature Survey and Design Study of Fume Cupboards and Fume-Dispersal Systems. Occupational Hygiene Monograph No. 4, 1980. Pp.82. Flexi ISBN 0-905927-50-8. (Science Reviews Ltd: 1980.) £7.50.

KAVANAGH, R. (ed.). Energy Systems Analysis. Commission of the European Communities. Proceedings of the International Conference held in Dublin, October 1979. Pp.678. ISBN 90-277-1111-9. (Reidel: 1980.) Dfl. 90, $47.50.

McAULIFFE, C. A. Hydrogen and Energy. Pp.109. ISBN 0-333-18432-7. (Macmillan: 1980.) £12.

MURR, L. E. (ed.). Solar Materials Science. Based on the Distringuished Lecture Series Sponsored by the New Mexico Joint Center for Materials Science, September–December 1979. Pp.788. ISBN 0-12-511160-6. (Academic: 1980.) $35.

ROBINSON, F. A. (ed.). Environmental Effects of Utilising More Coal. Special Publication No. 37. Proceedings of a Conference organised by The Council for Environmental Science and Engineering. Pp.204. Flexi ISBN 0-85186-805-3. (Royal Society of Chemistry: Burlington House, London, W1V OBN, 1980.) £9.50.

Earth Sciences

BALLANCE, P. F. and READING, H. G. Sedimentation in Oblique-slip Mobile Zones. Special Publication No.4 of the International Association of Sedimentologists. Pp.266. Flexi ISBN 0-632-00607-2. (Blackwell Scientific: 1980.) £14.50.

CARRERAS, J. *et al.* (eds). Sheer Zones in Rocks. Journal of Structural Geology Vol.2, Number 1/2. Pp.287. Flexi ISBN 0-08-026244-9. (Pergamon: 1980.) £15, $35.

DERRY, D. R. World of Geology and Mineral Deposits. Pp.110. ISBN 0-470-26996-0. (Wiley: 1980) £20.

GREESON, P. E., CLARK, J. R. and CLARK, J. E. (eds). Wetlands Functions and Values: The State of our Understanding. Proceedings of the National Symposium, Disneyworld Village, Lake Buena Vista, Florida, November 1978. Pp.674. (American Water Resources Association: St. Anthony Falls Hydraulic Laboratory, Mississipi River at 3rd Avenue, S.E. Minneapolis, Minnesota 55414, 1980) $49 + postage outside US and its territories.

JONES, D. K. C. (ed.) The Shaping of Southern England. Institute of British Geographers Special Publication, No.11. Pp.274. ISBN 0-12-388950-2. (Academic: 1980.) £14.60, $34.

MESSEL, H. *et al.* Surveys of Tidal River Systems in the Northern Territory of Australia and their Crocodile Populations. Monograph 9: Tidal Waterways of Castlereagh Bay and Hutchinson and Cadell Straits. Pp.132. Flexi ISBN 0-08-024801-2. (Pergamon: 1980.) £11.

SCHUMACHER, M. M. (ed.). Enhanced Recovery of Residual and Heavy Oils. 2nd Edn. Pp.378. ISBN 0-8155-0816-6. (Noyes Data Corporation: Park Ridge, New Jersey, U.S.A., 1980.) $48.

TURNER, F. J. Metamorphic Petrology: Mineralogical, Field, and Tectonic Aspects. 2nd Edn. Pp.524. ISBN 0-07-065501-4. (Hemisphere Publishing: 1981.) Np.

Biological Sciences

ALEXANDER, M. (ed.). Advances in Microbial Ecology, Vol.4. Pp.247. ISBN 0-306-40493-1. (Plenum: 1980.) $32.50.

APPLETON, D. R., SUNTER, J. P. and WATSON, A. J. (eds). Cell Proliferation in the Gastrointestinal Tract. Pp.428 ISBN 0-272-79597-6. (Pitman Medical: 1980.) £25.

BAKER, R. (ed.). Controlled Release of Bioactive Materials. Based on the Symposium at the 6th International Meeting in New Orleans, Louisiana, August 1979. Pp.473. ISBN 0-12-074450-3. (Academic: 1980.) $34.50.

BRITISH MUSEUM (NATURAL HISTORY) Man's Place in Evolution. Pp. 108. Hbk ISBN 0-521-23177-9; pbk ISBN 0-521-29849-0. (Cambridge University Press: 1980.) Hbk £12; pbk £3.95.

CARR, I. and DAEMS, W. T. (eds). The Reticuloendothelial System. A Comprehensive Treatise. Vol.1, Morphology. Pp.793. ISBN 0-306-40291-1. (Plenum: 1980.) $49.50.

COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY. Vol. XLIV: Viral Oncogenes. Parts 1 & 2. Pp.1322. ISBN 0-87969-043-6. (Cold Spring Harbor Laboratory: 1980.) Np.

DORNFELD, E. J. The Butterflies of Oregon. Pp.276. Flexi ISBN 0-917304-58-6. (Timber Press: P.O. Box 92, Forest Grove, Oregon 97116, 1980.) $24.95.

DURAND, J. R. and LEVÊQUE, C. Flore et Faune Aquatiques de l'Afrique. Sahelo-Soudanienne, Tome I. Collection Initiations-Documentations Techniques No.44. Pp.389. ISBN 2-7099-0521-3. (Editions de l'Office de la Recherche Scientifique et Techniques Outre-Mer: Paris, 1980.) Np.

FINN, C. A. Oxford Reviews of Reproductive Biology, Vol.2. Pp.273. ISBN 0-19-857535-1. (Clarendon Oxford University Press: 1980.) £20.

FUSTER, J. M. The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe. Pp.232. ISBN 0-89004-524-0. (Raven: 1980.) $22.

HINCHLIFFE, J. R. and JOHNSON, D. R. The Development of the Vertebrate Limb. An Approach through Experiment, Genetics and Evolution. Pp.268. 0-19-857552-1. (Clarendon/Oxford University Press: 1980.) £20.

JAMES, C. A. Manual of Assessment Keys for Plant Disease Pp.88. Loose leaf. (The American Phytopathological Society, 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) $10.

LOCKARD, J. S. and WARD, A. A. (eds). Epilepsy. A Window to Brian Mechanisms. Pp.286. ISBN 0-89004-499-6. (Raven: 1980.) $29.50.

McGANN, B. Behaviour, Health and Lifestyle: The Meaning of Human Health and the Essential Factors Which Enhance It. Pp.349. ISBN 0-906408-14-8. (Villa Books: Dublin, 1980.) £9.50 (UK), £10.45 (Eire).

MILNE, L. and MILNE, M. The Audubon Society Field Guide to North American Insects and Spiders. Pp. 989. Flexi ISBN 0-394-50763-0. (Knopf: 1980.) $9.95.

OTTOWAY, J. H. The Biochemistry of Pollution. The Institute of Biology's Studies in Biology, No. 123. Pp.58. Flexi ISBN 0-7131-2748-8. (Edward Arnold: 1980.) £2.50.

PANCHEN, A. L. (ed.). The Terrestrial Environment and the Origin of Land Vertebrates. The Systematics Association, Special Volume No.15. Proceedings of an International Symposium held at the University of Newcastle upon Tyne. Pp.634. ISBN 0-12-544780-9. (Academic: 1980.) £28, $87.50.

PIOMELLI, S. and YACHNIN, S. (eds). Current Topics in Haematology, Vol.3. Pp.271. ISBN 0-8451-0352-0. (Alan R. Liss: New York, 1980.) Np.

POLUNIN, O. Flowers of Greece and the Balkans: A Field Guide. Pp.568. ISBN 0-19-217626-9. (Oxford University Press: 1980.) £40.

PORTER, R. and WHELAN, F. (eds). Metabolic Activities of the Lung. Ciba Foundation Symposium 78. Pp.401. ISBN 90-219-4084-1. (Excerpta Medica: 1980.) $66.25, Dfl.136.

RATCLIFFE, D. The Peregrine Falcon. Pp.416. ISBN 0-931130-05-0. (Buteo Books: PO Box 481, Vermillion, SD 57069 USA, 1980.) $42.50.

RHOADS, D. C. and LUTZ, R. A. (eds). Skeletal Growth of Aquatic Organisms. Biological Records of Environmental Change: Topics in Geobiology, Vol.1. Pp.750. ISBN 0-306-40259-9. (Plenum: 1980.) $47.50.

ROTMAN, A. *et al.* (eds). Platelets: Cellular Response Mechanisms and their Biological Significance. Proceedings of an EMBO Workshop held at The Weizmann Institute of Science, Rehovot, Israel, April 1980. Pp.327. ISBN 0-471-27896-3. (Wiley: 1980.) £17.50.

SALANKI, J. and TURPAEV, T. M. (eds). Neurotransmitters: Comparative Aspects. Pp.514. ISBN 963-05-2601-8- (Akadémiai Kiadó, Publishing House of the Hungarian Academy of Sciences: H-1361 Budapest, P.O. Box 36, 1980.) $40.

SANER, G. Chromium in Nutrition and Disease. Current Topics in Nutrition and Disease, Vol.2. Pp.135. ISBN 0-8451-1601-0. (Alan R. Liss: New York, 1980.) $16.

SBARRA, A. J. and STRAUSS, R. R. (eds). The Reticuloendothelial System: A Comprehensive Treatise. Vol.2, Biochemistry and Metabolism. Pp.4342. ISBN 0-306-40292-0. (Plenum: 1980.) $45.

SCHAAD, N. W. (ed.). Laboratory Guide for Identification of Plant Pathogenic Bacteria. Pp.76. Loose leaf. ISBN 0-89054-028-4. (The American Phytopathological Society: 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) $10.

SIGEL, H. (ed.) Metal Ions in Biological Systems. Vol.11: Metal Complexes as Anticancer Agents. Pp.440. ISBN 0-8247-1004-5. (Marcel Dekker: 1980.) SwFr.115.

SKULACHEV, V. P. Biology Reviews (Physico-Chemical Aspects), Vol.1 (1980). Soviet Scientific Reviews, Section D. Pp.476. ISBN 3-7186-0020-A. (Harwood Academic Publishers: London, 1980.) Np.

SMITH, A. D., LLINAS R. and KOSTYUK, P. G. (eds). Commentaries in the Neurosciences. Pp.668. ISBN 0-08-025501-9. (Pergamon: 1980.) £30, $67.50.

STACE, C. A. Plant Taxonomy and Biosystematics. Pp.279. Flexi ISBN 0-7131-2802-X. (Edward Arnold: 1980.) £8.95.

STEIN, R. B. Nerve and Muscle: Membranes, Cells and Systems. Pp.265. ISBN 0-306-40512-1. (Plenum: 1980.) $18.95.

TAPPAN, H. The Paleobiology of Plant Protists. Pp.1028. ISBN 0-7167-1109-5. (W. H. Freeman: 1980.) $95.

TURNER, N. C. and KRAMER, P. J. (eds). Adaptation of Plants to Water and High Temperature Stress. Pp.482. ISBN 0-471-05372-4. (Wiley: 1980.) £21.50.

WARNER, N. L. (ed.). Contemporarty Topics in Immunobiology, Vol.11. Pp.299. ISBN 0-306-40419-2. (Plenum: 1980.) $32.50.

WEETE, J. D. Lipid Biochemistry of Fungi and Other Organisms. Pp.388. ISBN 0-306-40570-9. (Plenum: 1980.) $45.

WEINER, H., HOFER, M. A. and STUNKARD, A. J. (eds). Brain, Behaviour, and Bodily Disease. Pesearch Publications: Association for Research in Nervous and Mental Disease, Vol.59. Pp.388. ISBN 0-89004-480-5. (Raven: 1980.) $37.

WHATLEY, J. M. and WHATLEY, F. R. Light and Plant Life. The Institute of Biology's Studies in Biology, No. 124. Pp.91. Flexi ISBN 0-7131-2785-6. (Edward Arnold: 1980.) £3.40.

WHITAKER, J. O. Jr The Audubon Society Field Guide to North American Mammals. Pp.745. Flexi ISBN 0-394-50762-2. (Knopf: 1980.) $9.95.

WHITEMAN, P. C. Tropical Pasture Science. Pp.392. Hbk ISBN 0-19-859470-4; pbk ISBN 0-19-859471-2. (Oxford University Press: 1980.) Hbk £20; pbk £5.90.

WILDENTHAL, K. (ed.). Degradive Processes in Heart and Skeletal Muscle. Reserch Monographs in Cell and Tissue Physiology, Vol. 3. Pp.456. ISBN 0-444-80235-5. (Elsevier/North-Holland Biomedical: 1980.) $95, Dfl. 195.

WINFREE, A. T. The Geometry of Biological Time. Biomathematics, Vol. 8. Pp.530. ISBN 3-540-09373-7. (Springer-Verlag: 1980.) DM 59.50, $32.80.

ZENTMYER, G. A. *Phytophthora cinnamoni* and the Diseases It Causes. Monograph No.10. Pp.96. Pbk ISBN 0-89054-030-6. (The American Phytopathological Society: 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) $8.

Applied Biological Sciences

BRIDGES, J. W. and CHASSEAUD, L. F. (eds). Progress in Drug Metabolism, Vol.5. Pp.358. ISBN 0-471-27776-2. (Wiley: 1980.) £28.50.

COAKER, T. H. (ed.). Applied Biology, Vol. V. Pp.407. ISBN 0-12-040905-4. (Academic: 1980.) £24.50, $56.50.

DRAPER, H. H. (ed.). Advances in Nutritional Research, Vol.3. Pp.362. ISBN 0-306-40212-0. (Plenum: 1980.) $32.50.

DURIG, J. R. (ed.) Analytical Applications of FT-IR to Molecular and Biological Systems. Proceedings of the NATO ASI held at Florence, Italy, August 1980. NATO Advanced Study Institutes Series C: Mathematical and Physical Sciences 57. Pp.607. ISBN 90-277-1145-3. (Riedel: 1980.) Dfl.130, $68.50.

LEHNER, T. and CIMASONI, G. (eds). The Borderland between Caries and Periodontal Disease II. Proceedings of the 2nd European Symposium held in Geneva, Switzerland, February 1980. Pp.288. ISBN 0-12-792506-6. (Academic: 1980.) £18.20, $42.

NAIR, S. *et al.* (eds). Computers in Critical Care and Pulmonary Medicine. Proceedings of the 1st International Symposium, held at Norwalk Hospital, Connecticut, May 1979. Pp.427. ISBN 0-306-40449-4. (Plenum: 1980.) $39.50.

RAPADO, A., WATTS, R. W. E. and DeBRUYN, C. H. M. M. (eds). Purine Metabolism in Man–III: Clinical and Therapeutic Aspects. Advances in Experimental Medicine and Biology, Vol. 122B. Proceedings of the first half of the 3rd International Symposium, held in Madrid, June 1979. Pp.444. ISBN 0-306-40310-2. (Plenum: 1980.) $45.

RAUS, J., MARTENS, H. and LECLERCQ, G. (eds). Cytotoxic Estrogens in Hormone Receptive Tumors. Based on the Proceedings of a Workshop held at Diepenbeek, Belgium, September 1979. Pp.285. ISBN 0-12-583050-5. (Academic: 1980.) £14.80, $34.50.

SOMOGYI, J. C. and TARJÁN, R. (eds). Foreign Substances and Nutrition. Bibliotheca Nutritio et Dieta, No. 29. 16th Symposium of the Group of European Nutritionists, Budapest, November 1978. Pp.132. Flexi ISBN 3-8055-0621-X. (Karger: Basel, 1980.) Sw Fr. 83, DM 99, $49.75.

TWYCROSS, R. G. and VENTAFRIDDA, V. (eds). The Continuing Care of Terminal Cancer Patients. Proceedings of an International Seminar, Milan, October, 1979. Pp.280. ISBN 0-80-024943-4. (Pergamon: 1980.) £18, $40.

Psychology

ANDERSON, J. R. Cognitive Psychology and its Implications. Pp.503. ISBN 0-7167-1197-4. (W. H. Freeman: 1980.) £8.10.

MOWRER, O. H. (ed.). Psychology of Language and Learning. Pp.294. ISBN 0-306-40371-4. (Plenum: 1980.) $27.50.

RIEVER, R. W. (ed.). Applied Psycholinguistics and Mental Health. Applied Psycholinguistics and Communication Disorders. Pp.224. ISBN 0-306-40392-7. (Plenum: 1980.)

General

AHRENS, L. H. (ed.). Chemistry of the Moon. Physics and Chemistry of the Earth, Vol. 10. Pp.264. ISBN 0-08-020287-X. (Pergamon: 1980.) £43.

BAILEY, W. Man, Religion, and Science: A Functional View. Pp.242. Pbk ISBN 0-9604196-0-8. (Wm. Bailey: Santa Barbara, 1980.) $9.95.

BETTS, R. J. Business Economics for Engineers. Pp.246. Flexi ISBN 0-07-084112-8. (McGraw-Hill: 1980.) Np.

BOYLE, R. H. Bass. Pp.160. ISBN 0-393-01379-0. (W. W. Norton: 1980.) $27.50.

CALLAHAN, D. and BOK, S. (eds). Ethics Teaching in Higher Education. Pp.315. ISBN 0-306-40522-9. (Plenum: 1980.) $19.50.

COREA, G. Need for Change: Towards the New International Order. A Selection from Major Speeches and Reports with an Introduction. Pp.278. ISBN 0-08-02095-0. (Pergamon: 1980.) £12, $25.

HARTWIG, S. (ed.). Heavy Gas and Risk Assessment. Proceedings of a Symposium on Heavy Gas, Frankfurt/Main, September 1979. Pp.306. ISBN 90-277-1108-9. (Reidel: 1980.) Dfl. 70, $37.

MILUNSKY, A. and ANNAS, G. J. (eds). Genetics and the Law, II. Pp.480. ISBN 0-306-40388-9. (Plenum: 1980.) $29.50.

PERROT J. and VANDERMEERSCH, B. (eds). Paléorient, Vol. 5. 1979. Pp.310. Flexi ISBN 2-222-02613-X. (Édetions Du Centre National de la Recherche Scientifique: Paris. 1980.) Np.

PRASAD, E. A. V. Ground Water in Varahamihira's Brhat Samhita. Pp.351. (E. A. V. Prasad, Sri Venkateswara University, Tirupati, India: 1980.) Rs. 125.

ROBINS, E. Secret Eden: Africa's Enchanted Wilderness. Pp.128. ISBN 0-241-10423-8. (Elm Tree: London, 1980.) £8.95.

SCHEINER, A. P. and ABROMS, I. F. The Practical Management of the Developmentally Disabled Child. Pp.461. ISBN 0-8016-0061-8. (C. V. Mosby: 1980.) Np.

nature

5 February 1981

15/2/82

Two views of the causes of cancers

What is the cause of cancer? This question has for the past half-century engaged physicians, biologists and laymen. It is, of course misstated. There is no one kind of cancer and no single cause. Readers of *Nature* will, however, have been struck by the contrast between the two substantial articles on the subject published in the past few weeks. Three weeks ago (*Nature* **289**, 127–130; 1981), Epstein and Swartz, under the title "Fallacies of lifestyle cancer theories", set out to demolish the earlier arguments of Peto (*Nature* **284**, 297; 1980), who had himself complained about Epstein's book *The Politics of Cancer.* The issue in dispute is the relative weight to be given to voluntary and involuntary sources of chemical pollution in the causation of human cancer, but there are also subsidiary disputes about questions such as the proper interpretation of the epidemiological data. Is there now an epidemic of cancer, or is it merely that more people are dying of cancer because there are fewer competing (or masking) causes of death? The second article, by Cairns (*Nature* **289**, 353–357), tackled almost the same question in a very different way. His objective was not to demonstrate causes but to point out the weaknesses of the mechanisms suggested in recent years as plausible explanations of how some kinds of cancers are caused. The truth, says Cairns, cannot be as simple as we would like to think. And much of what is said about the origin of cancer must be wrong.

How can these very different points of view and styles of argument be reconciled? Does it matter if they are not? And how should people, not simply specialists, approach the question of what causes cancer? Cairns's argument deserves the most careful reading, chiefly because it shows how intricate the question is. The biology of cancer is not fully explained by the now familiar model in which (the story goes) damage to a DNA strand in the chromosome of some somatic cell can, if the damage is appropriately placed, give rise to a clone of cancerous cells. The model has the virtue that it accommodates very different carcinogens — particle radiation and chemical carcinogens, for example. Fast particles can obviously damage DNA molecules physically. The metabolites of at least some chemical carcinogens include materials likely to react with DNA, with consequences that are plausibly similar. So such events may be first causes, in Hume's sense, of some kinds of cancer.

The model is useful, and should not be discarded. But it is also incomplete. It does not, for example, account for the length of the induction period for most kinds of cancer. If damage at a single point along a DNA strand can be a sufficient cause of cancer, why does not the cell thus affected instantly become the founding member of a cancerous clone of cells? And why is it that (as Cairns points out) such chromosomes of cancerous cells as can be seen are most often grossly abnormal in ways indicative of drastic rearrangement, not of a defect at a single point? The simple model of cancer causation can be elaborated to accommodate uncomfortable observations such as these. Conceivably, the induction period and the eventual appearance of gross chromosome abnormalities are both consequences of the need that a potentially cancerous clone of cells should evolve into a form no longer subject to the unknown restraints that normally inhibit competition between tissues within the same organism. Damaging a single point along the length of a DNA molecule may still be a sufficient cause of cancer. The chances are, however, that it is not. How else is it possible to account for the well-founded evidence that people differ genetically in their susceptibility to some kinds of cancer? And what, in any case, is to be made of Cairns's argument about the inherited disease xeroderma pigmentosum? The underlying trouble is a defect of the normal biochemical mechanism for repairing damaged DNA molecules. The observation that people afflicted with this genetic disease are prone to skin cancer and melanoma is consistent with the simple model. Ultraviolet light is a well known cause of damage to DNA molecules. But why are not the same people also more susceptible to other kinds of cancer caused by other carcinogenic agents? Cairns asks the question. Nobody knows the answer.

The implication of Cairns's arguments is clear and also depressingly familiar. The phenomenon of human cancer is a problem in biology which is not yet understood. It will be time enough to talk about causes when mechanisms have been worked out. That will require not merely more of the molecular biology and cell biology now supported by the cancer research agencies but also a better understanding of the processes of differentiation, the physiology of homeostasis and even of some long-standing issues in old-fashioned biology. The prospect is nevertheless far from hopeless and, as in other fields, there is no reason why the development of better means of treating cancer should wait on a full understanding of the mechanisms. But, for the time being, the causes of cancer in Cairns's sense are obscure.

How then can Epstein and Swartz pretend such confidence about the causes of cancer? Their starting point is that a cancer requires exposure to some carcinogenic agent, perhaps a chemical in the environment or some component of tobacco smoke. Their methods are epidemiological but not necessarily suspect on that account. Statistical analysis can establish relationships between causes and effects. Epidemiology has, for example, demonstrated the occupational hazards of vinyl chloride monomer. It is also now to be taken as a fact that people who smoke cigarettes are more likely than others to develop lung cancer. The habit of smoking cigarettes is thus in some sense a cause of lung cancer. The link is clear enough for individuals prudently not to smoke cigarettes and for governments to encourage abstention. Yet the habit of smoking is not a first cause of lung cancer. That distinction no doubt lies with a group of chemicals not yet conclusively demonstrated. The mechanism by which lung cancer develops is not known, with the result that there is even room for argument about the influence of unknown genetic factors in the development of the disease. Such difficulties do not of course vitiate the epidemiological approach to cancer. On the contrary, there is undoubtedly great benefit to be won, both in the recognition of specific hazards and in the wider understanding of the natural history of cancer. Prudent practitioners are, however, careful not to make too much of the associations they find between effects and their supposed causes.

Epstein and Swartz appear to be more than a little cavalier in their regard for that simple rule. The trouble seems to start from *The Politics of Cancer,* whose purpose was to argue that governments (but especially the United States government) should do more to regulate occupational and environmental carcinogens, and which has set the tone for much of the discussion that has followed. The purpose, of course, is entirely legitimate. It is the method that is suspect. Epstein, writing for a general audience, has assembled an impressive list of case studies of chemicals whose use has been implicated in cancer, accurately or otherwise. The more gentle among his readers have no doubt by now been persuaded that there is no hope of avoiding cancer while they continue to live in advanced societies such as America's. Some of them, perhaps for other reasons, will no doubt have

0028-0836/81/060431-02$01.00

embraced the conclusion that the chemical industry, by the scale of its recently increased manufacture of organic chemicals, has become the chief hazard. Yet many of the items in Epstein's indictment are, to say the least, disputable — the carcinogenicity of saccharin, for example, as Peto quite properly pointed out nearly a year ago. Epstein's comment on this point (*Nature* **289**, 115; 1981) implicitly confirms that view. The nature of his conclusion, and the clarity with which it is stated, seem entirely at odds with the simple rule that epidemiologists should acknowledge the possibility that their statistical association may be wrong. Self-doubt is not merely seemly but persuasive.

But is not an appearance of self-doubt inappropriate in a book whose purpose is polemical? One objective of *The Politics of Cancer* is after all to persuade governments to change their ways. Epstein, who appears from his bibliography to be a frequent witness before congressional committees, does not lack opportunities for putting his point of view. On this occasion, however, the result may have been so to scare such a large number of readers of the book that they in turn will put pressure on their congressmen, forcing a change of policy towards the chemical industry and its products born not of reason but of fright. There can be no circumstances in which it is permissible for professional scientists to hazard the reputations of their subjects in such a cause. The overstatement of a plausible case is not permissible. To conduct scientific arguments in immoderate language is bad for the public reputation of science. This is why Cairns's way of looking at the problem, with all its inconclusiveness, is so much to be welcomed.

University empire confronts a crisis

What on earth is to be done about the University of London, which is not so much a university as a relic of a university empire? It is a huge and ungovernable federation of more than forty titularly autonomous institutions, more than half of them concerned with medical education. There are eight non-medical colleges providing a general undergraduate curriculum and a further ten with more specialized interests, some only at the graduate level. The schools differ greatly in size, strength and reputation. Some, such as the London School of Economics and Imperial College, are international centres of scholarship — and Imperial College is financed separately from the rest of the university on the grounds of its national importance. Others, not always through their own fault, are at the other end of the spectrum of academic excellence. In 1979–80, there were 30,000 undergraduates and 10,000 graduate students, 18 per cent of them from overseas. London's financial prospects are thus bleaker than those of other British universities because its proportion of overseas students is greater; it stands to be hit harder by the government's insistence that overseas students must now be charged full economic fees (see *Nature*, 29 January). The university has long since abandoned its role as a provider of school-leaving examinations for most of the British Commonwealth and has even, more recently, given up the invidious role of providing other British institutions of higher education with certificates of good housekeeping. Now, like the British Empire before it, the university is faced with the prospect of collapse or, at least, of rapid contraction.

How is the problem being tackled? London's first step was predictable — setting up a committee. There is the obligatory outside chairman (Sir Peter Swinnerton-Dyer, vice-chancellor of the University of Cambridge) and marching orders to suggest how resources should be redistributed so as to keep the university in being ("academic excellence" etc.) at less cost ("having regard to the need to make financial economies" etc.). The committee, which is concerned only with the non-medical parts of the university, has sought to avoid the ructions that greeted the report of the Flowers committee on medical teaching last year by publishing last week an interim discussion document. The result is a collection of figures for the costs of teaching students in different subjects in the various colleges, a solemn warning that the years ahead will see income fall by 15 per cent and student numbers by 10 per cent, and hardly a hint of the recipe for reorganization the committee has promised by the end of the year. The document is thus not so much an interim report as an incitement of the lobbying and horse-dealing that will occupy the University of London for months to come.

That something, almost anything, must be done is crystal clear. For decades, the university has been almost as much concerned with attempts to reform itself as with the teaching of students. The practical difficulty is that change is usually unacceptable to the institutions most affected, which within the federal structure of the university have the final say. The Swinnerton-Dyer committee will succeed where others have failed only if it can persuade the constituent institutions of the university to bury their separate interests for the sake of the university to which they belong. The weakness of the interim document is that it seems designed to frighten the separate colleges into compliance with some unknown grand design by the threat of penury. The document thus urges that the university and its colleges should urgently consider how to manage redundancies among academic and other members of staff: tenure is plainly no longer sacred. It makes little of the threat, uppermost in the minds of many colleges in the past few months, that some of them may be closed or merged with others; there is merely a statement that the university can distribute as it thinks fit the funds it receives from the University Grants Committee (subject only to what the senate and the court of the university have to say). The strongest hint of what is to come is that university teachers might with advantage teach in more than one college. These possibilities, sensible enough, do not amount to a recipe for a stronger university. And they smack more of the stick than the carrot.

For the sake of the university, the Swinnerton-Dyer committee must find a more positive case to put a year from now. Although the interim document is full mostly of figures, it does acknowledge that in the end academic considerations will be paramount. What should these be? The committee must first count its blessings, which are various but not always obvious. The larger general colleges are comparable in size as well as quality with independent universities elsewhere in Britain. Some of the smaller colleges have a claim on public attention (and funds) for their tradition of educating less highly qualified entrants. Birkbeck College is the only institution of its kind in Britain with the objective of providing part-time students with an undergraduate education. Five of the independent schools (architecture, agriculture, education, pharmacy and veterinary science) are predominantly vocational; their place within the university calls for a frank but sympathetic appraisal. Should they all continue and, if so, should they continue to be supported from central funds as if their students were ordinary students? The risk is that the committee will be mesmerized by the reputations of the outstanding schools, letting the others go hang.

Problems of organization are also likely to obtrude. What kind of university should London become? The paradox in the past decade is that all parts of it have become more independent of the centre, which suits the larger schools well enough but makes it harder for the others to enjoy a sense of being part of a university. Nobody wants to see a revival of the dominance of bureaucracy, but the question must arise whether it is not the proper function of a university such as that in London to be a nursery for new kinds of university institutions — and whether the time has not arrived when places such as Imperial College should be encouraged to go their own way, to become autonomous in name as well as in fact. As things are, the equitable allocation of resources is unmanageable. How can the giants be denied? And how then can the pygmies enjoy a sense of partnership in a common enterprise — not that of being mere pygmies?

0028-0836/81/060432-01$01.00

Cutback on West German development

Money now in short supply in Bonn

The West German research ministry is being forced by economic problems to reassess priorities. Last week, the Bundestag began a discussion of the proposals put forward by the new science minister, Andreas von Bulow, whose plan for 1981 will increase the budget by only 5.2 per cent, now less than the rate of inflation. This new small increase has come as a shock to West German planners who are unused to having their budget cut.

The increase is less than had been forecast in the three-year forward plan published a year ago, which promised an increase of 10 per cent (in money terms) for 1981. The new research budget is not, however, out of line with revisions of other government expenditures, consequences of the worsening economy.

The research ministry, which controls the lion's share of funds for German research, is planning to save money by cancelling some new projects. Mr von Bulow says that he does not want to delay projects already under way. Expensive demonstration projects are likely to suffer most, but the government hopes that some of their cost can be taken over by the utilities (in nuclear power and coal), by industry, the Länder or local government. University research may do better because of the commitment to keeping university researchers in employment. The aim is to keep the proportion of the budget spent on basic research at just over 30 per cent.

Three particular projects are affected. A demonstation track for testing an advanced passenger train and a new tokamak device for fusion experiments at the Institute for Plasma Physics in Garching have already been abandoned. The cancellation of the demonstration track is expected to save DM30–40 million this year and DM 250 million by the end of 1983. The fusion device, known as Zephyr, would have cost an estimated DM500 million to build at 1980 prices. Its cancellation has removed a major project from the work of the plasma physics institute and has led the ministry to ask the institute to become more involved in the Joint European Torus (JET) under construction in Britian.

A decision has yet to be reached on the fate of the third project, a coal liquefaction plant being built in Virginia, United States. Germany and Japan are each expected to contribute 25 per cent of its cost and the Bundestag is having to assess the repercussions of cancellation.

Research fields likely to receive more than a 5.2 per cent increase include coal, radioactive waste disposal, microelectronics, biotechnology and polar research. The likely losers include transport, building, space research and the large nuclear research centres. University building is also to be reduced and savings will be made at the Deutsches Electronen-Synchrotron (DESY) by reducing the operating time of PETRA, the electron-positron machine. In all fields, the aim is to cut spending on development in favour of research.

The universities will nevertheless be affected to some extent. The ministry helps to support some specific university projects even though most of the money comes from Länder and from the Deutsches Forchungsgemeinschaft, the chief source of funds for university research. The Länder, which support the basic infrastructure including staff salaries, are also cutting increases in their budgets, but they are limited by their commitment to keep tenured staff in employment. The Forschungsgemeinschaft's budget is intended to increase from DM780 million in 1980 to DM830 million in 1981, an increase of 6.4 per cent.

In a revision of the forward plan to 1983, savings of DM2,000 million on the entire federal government research budget are planned for the next three years. The Bundestag's deliberations on the research budget will last for another few months, but a radical departure from the proposals is unlikely.

Judy Redfearn

Polish students sit in at universities

Poland now also has its student sit-ins. At the University of Lodz, 6,000 students drawn from the university, the technical university and the medical school have so far been involved in a sit-in which began in mid-January. The academic work of these establishments is virtually paralysed. The students' demand is for university autonomy and a less rigid curriculum.

Lodz is not unique. At the end of November, there was a sit-in lasting several days at the University of Warsaw, while in January some 200 final-year students of the Wroclaw agricultural academy held a sit-in over issues such as the system of examinations and credits. In particular, they demanded that the present system of rigid lectures and examinations in "socio-political sciences" should be replaced by informal seminars.

The issue of "socio-political" teaching is one of the fifty-odd grievances of the Lodz protesters. The range of their complaints is considerable and, as the Minister of Science, Higher Education and Technology, Janusz Gorski, said after talks with the university senate last week, many of them are impractical. A maximum of 3 months military service for students, for example, would be meaningless. The most striking feature of the Lodz protest, however, is not the range of demands but the range of participants; the action is backed not only by the new Independent Students' Association and the Solidarity chapters of the academic staff, but also by the university party committee and the "old" organizations — the Association of Polish Socialist Students and the Union of Polish Teachers.

The Lodz protest is a symptom of discontent that runs throughout all Polish academic life. Implementation of the Gdansk Accords, signed at the end of the summer, is slow, and delays in drafting new bills on academic autonomy and the reduction of censorship to the minimum needs of national security have evoked much anxiety that the terms of the accords may remain paper pledges only.

Thus the censorship bill, promised for 30 November, which is to take academic publishing out of the censors' control, has not yet appeared, even though the extended deadline of 15 January has been passed. Two draft bills have in fact been produced — one by the Minister of Justice, Jerzy Bafia, and one by a "public drafting committee", headed by Warsaw philosopher Dr Klemens Szaniawski. The problem of reconciling the two has become bogged down in the question of accountability. Minister Bafia would like the Office of Press, Publications and Performances (the censorship) to remain, as now, accountable to the Council of Ministers (the Cabinet). Dr Szaniawski's committee, however, wants it to be answerable to the Sejm (Parliament). Many of the proposals for "social renewal" put forward in recent months have called for the Sejm to play a more active role in the political life of the country.

In the absence of legislation, some universities have been quietly making their own reforms. Some rectors have made available to their students the university holdings of "restricted" texts.

Anomalies inevitably abound. Thus last week the Procurator General's office warned that issuing uncensored publications still carries a potentially heavy gaol sentence. Yet a few days previously, the censors had passed an issue of the Catholic weekly *Tygodnik Powszechny* which described lectures given in Gdansk by the clandestine "Society for Academic Courses" (Flying University) — lectures which, in pamphlet form, are regularly published by the underground press.

Censorship is only one of the burning academic issues. Dr Aleksander Gieysztor, the newly elected president of the Polish Academy of Sciences, sees the

0028-0836/81/060433-02$01.00

democratization of academic life as a two-stage process. The first, minor, stage, he observed recently, is "already under way" —"the introduction of regulations leading to "autonomy, democratization and better information". The second, and more important stage, is to reform the "overdeveloped" administration, and to ensure that scientific committees truly represent the interests of their members.

Cutbacks in the scientific bureaucracy will be greatly welcomed in the natural and technical sciences where the hierarchical classification of research projects, introduced in the early 1970s, means not only that virtually all research has to be orientated to some need of the national economy but also that there should be a vast bureaucracy to deal with funds and administration.

Given less paper-shuffling and a certain latitude in undirected research, few academics would overlook the needs of the economy. On the contrary, Solidarity has already set up its own "All-Poland Coordinating Committee for Science" and the Solidarity chapters in the various universities have put forward their own proposals for what should be done.

Some proposals are ambitious, such as that drawn up two weeks ago by the Jagellonian University of Krakow which calls for action on censorship, visas for academic travel, health service reform, pollution and the amelioration of the social conditions leading to alcoholism. Others are more specific, such as the defence issued by the Maria Curie-Sklodowska University of Lublin of the claim of Poland's 3.5 million peasant farmers to the right to form a trade union. **Vera Rich**

Telecommunications suit

Pact postponed

For lack of an assistant attorney-general, the Reagan Administration has had to postpone the settlement of its anti-trust suit against AT&T, the telephone company. Details of the settlement, which were to have been made public last Friday, will not now become known until 2 March, the date set by the trial judge, Howard H. Greene, for the submission of a settlement between the telephone company and the US Department of Justice.

The suit was begun in November 1974, and the trial opened on 16 January with formal statements by the two parties, followed by a recess. The further postponement is at the request of the Department of Justice, which explained in a letter to the court that it could not proceed until the Administration had appointed an assistant attorney-general with responsibility for anti-trust affairs.

The settlement, the essence of which has already been agreed, will have an important influence not merely on the future shape of the telecommunications industry in the United States but also on the future role of the Bell Laboratories within it. The terms of the settlement are almost certain to include AT&T's separation of its ownership and management of the United States telecommunications network from its manufacturing subsidiary, Western Electric.

In a letter to the court supporting the request for a further postponement of the settlement, AT&T said last week that it had already bound itself to certain undisclosed conditions. There is increasing speculation that, when published on 2 March, these will be found to embody the reorganization of the corporation announced in August last year, involving separate management and accounting for the operating and manufacturing arms of the business. The future of Bell Labs, undecided then, is still in the air.

There is a sense in which the anti-trust suit by the Department of Justice has been overtaken by events, especially by the ruling of the Federal Communications Commission (see *Nature,* 11 September 1980) that companies wishing to supply terminal equipment cannot also manage the common carrier networks that connect them. This decision of what is known as the "Second Computer Inquiry" seems to have been the spur for the reorganization announced by AT&T last summer. Separately from the anti-trust suit against it, AT&T is applying to the courts to have the deadline for the reorganization decreed by the commission (1 March 1982) extended on the grounds that it cannot bring about the "biggest corporate reorganization in history" by then.

European Community

Squeezing grants

Brussels

Haggling among member states is likely to force down the increase in the budget proposed by the European Community for its scientific and technical training programme. Last Thursday, the member states' permanent representatives to the Community met for the second time and finally decided to allocate to it 8.8 million EUA (European Units of Account) or £4.5 million, which is 3 million EUA less than the original modest proposal. A final decision has still to be taken by the Council of Ministers, but this is expected within the next two weeks. Depending on how the European Community chooses to allocate grants, there will probably be the equivalent of about 340 scholarships.

While the United Kingdom was in favour of keeping the programme at its previous size of 4.5 million EUA, the French were insistent that it should be reduced. Such was their concern that the French felt unable to give their consent without calling Paris for permission. Other member countries were more worried about the number and size of the individual grants. Ireland was in favour of smaller grants for more researchers while the Danes preferred the idea of devoting larger sums to contract researchers of a high quality.

The four-year programme is intended to give students and junior lecturers a chance to study in another part of the Community. This can be either in one of the joint research centres (Geel, Ispra, Karlsruhe or Petten) or in research institutions in another member state with whom the Commission has concluded research contracts. Those wishing to study at home will be forced to accept a cut in their grant of 25 per cent. About 320 grants were allocated in the last programme but since the new programme must include grants to Greek researchers, there will be no increase in the number available in the rest of the Community.

Normally, grants are awarded for up to a year, but they may be extended to two years for those undertaking major research projects. Part of the aim of the programme, however, is to promote contracts between those responsible for the implementation of research programmes and the relevant education and training establishments. Details can be had from the Directorate-General for Research, Science and Education, Rue de la Loi, Brussels 1049.

Last week's meeting also decided on the budget for the next four-year research programme on the environment. The Commission's initial proposal of 51 million EUA was reduced to 42 million EUA. The programme includes the continuation of a major effort in the field of climatology and other areas covered are the treatment of sewerage sludge, air pollution and organic micropollutants in water. Here again, France was keen to see the budget cut, in this case to 35 million EUA. The Greek delegate spoke for seismological studies, which won agreement in principle. Formal research programmes are at present being examined in the Council of Ministers. **Jasper Becker**

UN energy conference

Prospects improve

Washington

Prospects are looking up for the forthcoming United Nations Conference on New and Renewable Sources of Energy, due to take place in Nairobi, Kenya, in August.

Last summer, the second meeting of the conference's preparatory committee found the secretariat's efforts so far behind schedule that postponement was seriously considered. In addition, the Kenyan government began to develop cold feet about the costs of staging a major international conference, the first since the 1979 Vienna meeting on science and technology for development, and to mutter about withdrawing its invitation.

But neither threat has materialized. And when a group meets in New York at the end of this month to process the reports of eight

0028-0836/81/060434-02$01.00

technical panels on the prospects for different energy sources, it is expected to be in a spirit of optimism that few thought possible six months ago.

One reason is the recent appointment of a new secretary general for the conference, Mr Enrique Iglesias, a Uruguayan who is also executive secretary of the Economic Commission for Latin America. Mr Iglesias, thought to be a candidate for Mr Kurt Waldheim's job as Secretary General of the United Nations, succeeds Mr Mohamed H. Gherab, who has recently been charged with accepting loans from subordinate officials.

A second reason for guarded optimism about the energy conference is that at present the UN secretariat seems to be successfully treading the delicate line between the technical and the political.

It has now been generally accepted that, given a lack of time and resources, the conference can only achieve a limited "state of the art" review of new and renewable energy sources (including solar, geothermal, wind, tidal, biomass, oil shale and hydropower energy). The main focus is therefore likely to be on the institutional mechanisms that can accelerate research and development on these various energy sources — and the obstacles that stand in the way of their implementation.

It is unlikely, however, that any new institution will emerge from the Nairobi conference. Proposals are more likely to be along the lines of a scheme being worked on by the World Bank to form an international network of energy research centres, with a clearly defined set of global priorities distributed in a way that minimizes duplication of research efforts.

Parallel initiatives are also under way inside the United Nations Development Programme (UNDP), which last summer set up a new energy account. UNDP has already received $3.5 million from the World Bank to carry out a survey of the energy needs of 60 developing countries, as part of the ambitious scheme announced by bank president Robert McNamara last year for a programme of energy loans and investment to total about $25,000 million by 1985.

The main concern of the UN conference organizers is devize a set of policy proposals that will be sufficiently specific to meet the recommendations of the technical panels, but sufficiently broad to generate the necessary political support.

One idea under discussion, for example, is a coordinated effort to replant trees that have been cut down for fuel. The World Bank has already proposed raising $1,000 million towards such a scheme, on the basis that a comparable sum would be found by the individual countries concerned.

Inevitably, there will be points of conflict, some of which have already come to the surface. Little attention, for example, will be given to the environmental efforts of different energy sources, a problem at present of greater concern to the developed than the developing world. Similarly, the United Nations General Assembly has explicitly stated that the conference will not consider conservation technologies, even though many in the developed countries feel that a reduction in demand is one of the likeliest ways of tackling the energy problem.

There is also dispute about the role of the oil-producing nations. These hold many of the important cards, both in terms of their ability to affect energy prices and in having the cash surpluses available to which access would be needed for any major investment schemes. Their attitude towards the conference remains ambiguous, although Mr McNamara has made it clear that the success of his proposals depends largely on their support.

Given the actual and potential disagreements on each of these topics, therefore, the success of the Nairobi conference still hangs in the balance. But some see light at the end of the tunnel, and argue that although time is getting short, elements such as the World Bank and UNDP initiatives, existing trends in foreign aid budgets and the completion of the technical reports are sufficient to allow a successful outcome. **David Dickson**

Ariane space launcher

Still in trouble

The problems of Ariane, Europe's hope for a space launcher, are not over yet. The third test flight, delayed after a failure in a first stage engine on the second test flight last summer, is now unlikely to get off the ground until well after June, the date to which the European Space Agency is still clinging. The difficulty is that no precise explanation of what went wrong with the second flight has yet emerged, and correcting the fault still seems to be a matter of trial and error.

A new schedule for the Ariane programme will not be released until one of two modified fuel injection systems identified as the cause of the fault has been chosen after tests expected to last 4–8 weeks.

The second test flight failed after high amplitude oscillations at 2,300 and 2,700 Hz developed in one of the first stage engines. Tests last October with modified fuel injectors of improved tolerance seemed to solve the problem at 2,300 Hz, but the 2,700 Hz oscillation remained.

Oscillations at these frequencies had not shown up in early tests of the Ariane engines and injectors. But tests up to the middle of last month indicate that the injectors used in early Ariane development differ from new ones, suggesting that they have been modified during assembly, preparation or testing. Nobody has been able to discover precisely how the injectors were modified, but the space agency says it is now trying to increase the margin of error acceptable in the injector design.

The latest line of attack is to test two modified injectors in the hope that one of them will be free from oscillation problems. Officials are hopeful that one system can be chosen within the next couple of months and work can begin on preparing the third test flight. Even if there are no more setbacks, however, a launch in June seems optimistic.

Meanwhile, the French space agency, whose idea it was that Europe should build its own space launcher, is now planning to propose another multi-million dollar venture to the European Space Agency. France wants to be in on the "industrialization" of space in the 1990s and hopes to suggest a remote controlled space laboratory, along the lines of the Russian Salyut space station.

The plan is to build a laboratory for materials processing under micro-gravity that would be placed in geosynchronous orbit by Ariane. The laboratory would be serviced by an expendable vehicle that would deliver supplies and bring processed materials back to earth. A third spacecraft, also in orbit, would be capable of building large structures and experiments for use by the laboratory. The proposal is still only a feasibility study, and will not be put to the space agency before the end of the year. Before then, the French plan to sound out other European nations; preliminary discussions with Germany have already taken place. **Judy Redfearn**

UK research councils

Getting off lightly

Next year's allocation of the "science vote" of the UK Department of Education and Science is being accepted thankfully by the UK research councils. The thinking may be that if Prime Minister Margaret Thatcher's government can make £200 million cuts in the Conservatives' defence budget, the councils are lucky not to suffer even more.

The Science Research Council, for example, gets £174 million for 1981–82 (at October 1979 prices). Converted to average prices for the current year (say October 1980) this works out at £198 million, compared with current spending for 1980–81 of £204 million — a 3 per cent decrease in real terms. In its annual report for 1979–80, published last November, the council expected "a modest increase" in the next two years; yet a spokesman said this week that the council was "pleased" that the government was treating science and engineering so well.

The Science Research Council's nominal spending on science within the United Kingdom is at present £150.5 million (for 1980–81), and on international subscriptions a nominal £44.5 million. But the latter was calculated at 1979 exchanges rates, since when the pound has strengthened sufficiently to reduce the subscriptions bill by some £9 million to

0028-0836/81/060435-02$01.00

around £36 million. The council has been allowed to use some of this saving to offset its overexpenditure on domestic grants this year (see *Nature* 13 November). Next year, the Science Research Council projects a nominal expenditure of £158.5 million on UK-based science, with £39.5 million on subscriptions (at the same October 1980 domestic prices and exchange rates). Since the pound is unlikely to strengthen further, at least to the degree it strengthened in 1979–80, and the grants to which the council committed itself in 1980 will be continuing, 1981–82 seems set to be a difficult year for SRC.

Among other councils, the Agricultural Research Council has registered a 3 per cent rise from £29.5 million to £30.5 million, and the Natural Environment Research Council remains virtually static at £39.4 million. The Medical Research Council gets £73.2 million, including £13.9 million contract research money recently transferred from the control of the Departments of Health and Social Security. The Social Science Research Council will receive £16.1 million, the British Museum (Natural History) £6.5 million, and the Royal Society £3.3 million. These figures will be adjusted in October this year to account for the inflation rate.

Robert Walgate

Canadian research

Aiming high

Washington

Canada's Liberal Government brought its election promises a step nearer reality last month when it committed itself to increasing the proportion of the country's gross national product spent on research and development from 0.9 per cent to 1.5 per cent by 1985. The commitment is ambitious. Even though the federal government expects a large part of the burden to be carried by the private sector encouraged by tax incentives and other inducements, its own contribution will have to rise from $1,000 million to $2,600 million in the next five years — in real terms a growth of 8 per cent a year, allowing for inflation.

Much will also depend on the attitude of the provinces, which through their support of universities contribute about 20 per cent of the nation's effort in research and development. Relationships between the federal and provincial capitals have been cool recently, but officials in Ottawa hope that the need to spur the competitiveness of Canadian industry is one area in which it may be possible to reach agreement.

It will be some time before the broad commitments, announced in Toronto by the Minister of State for Science and Technology, Mr John Roberts, are translated into specific budgets for the various research agencies. And the scientific community is therefore reserving judgement until it sees the colour of the government's money. But for now, the minister's announcement can be taken as a sign of good faith, since if nothing else it provides a broad framework within which sectoral policies can be worked out.

Spending on research and development has become something of a political football in Canada in recent years, to the surprise of foreign observers who frequently point out that — as the United Kingdom discovered the early 1970s — research spending can be a misleading indicator of industrial strength. Nevertheless, the Canadian public has been convinced that one of the reasons for the country's dependence on foreign investment capital, particularly from the United States, has been its relative lack of an indigenous technological capability. And both the Liberal and Conservative parties have competed with bids to raise research spending to correct the situation.

During last year's election campaign, the Liberals had promised yet again to reach the 1.5 per cent figure, but with no specific target date. Now they have fixed on 1985, and determined that by this time 50 per cent of the nation's research and development spending (equivalent to 0.75 per cent of its gross national product) will come from industry, with a third from the federal government and the rest from the provinces and universities.

Reaching these targets will require a considerable shift in responsibility from the public to the private sector. Industry is being asked to increase its share of the budget from 40 to 50 per cent, while the federal government's contribution would fall in relative terms from 40 to 33 per cent, even though the actual amount of money would increase.

To meet these targets, industry would have to increase its spending by 27 per cent a year — a goal which has already been viewed with a certain amount of scepticism, since the government will have to devise an appropriate set of indirect mechanisms to encourage such growth. Officials in Ottawa are now carrying out a quantitative analysis of the effects of previous policy measures, such as tax breaks on research investment and industrial training grants. One official admitted last week that the studies have so far not been too productive; but there is optimism that some fine tuning can be achieved.

Canada's efforts in this direction bear a close resemblance to the recent domestic policy review of industrial innovation carried out two years ago in the United States under President Carter.

One significant difference, however, is that the Canadian government supports a strongly sectoral approach, identifying areas of high technology — such as nuclear energy, telecommunications or space technology — where it feels that Canadian industry has a particular role to play on international markets.

Although there were structural elements in the policy recommendations put forward by President Carter eighteen months ago, these were overshadowed by broader proposals aimed at encouraging the general climate for innovation. And with the election of President Reagan, even schemes such as the joint government/industry sponsored Cooperative Automotive Research Project, a pet project of the former president's science adviser, Dr Frank Press, are under threat from free marketeers.

In Canada, the sectoral approach has been strongly supported by groups such as the Science Council, whose chairman, Dr Claude Fortier, wrote in the council's recent annual report that government action is urgently needed along sectoral lines to stimulate innovation and productivity.

This concern is reflected in the government's new proposals. Of the new money which it is proposing to make available for the support of research and development, one half will go to support work in the private sector in clearly defined categories, one third to the government's own mission-oriented laboratories and the rest to universities and other research institutes.

But how much Mr Roberts' announcement indicates a genuine commitment by the Canadian government, already facing a growing budget deficit to increased funding, and how much is merely a political exercise, has become a subject of intense debate.

David Dickson

Environmental cadmium

Europe to ban?

Brussels

Moves to restrict the use of cadmium are gathering momentum in Europe. In Brussels, the European Commission's directive laying down limits for the discharge of cadmium into the aquatic environment is close to being adopted, before it goes to the Council of Ministers, while in Germany, the Home Office Minister has caused a furore in industrial circles by announcing that a new report on cadmium may lead to tough restrictions on its use.

This closely follows news of a Swedish ban due to come into force in July 1983, which has caused much unease among European producers of plastics, stabilizers and pigments, many of which contain cadmium compounds. As a result of research carried out by the Danish National Agency of Environmental Protection, the EEC Council of Environment Ministers has concluded that present environmental levels of cadmium are potentially harmful, the cadmium accumulating in the lungs, bone tissue, and particularly in the kidneys. Cadmium accumulates slowly in man, reaching dangerous levels at around the age of 50. Heavy smokers are thought to be most at risk.

German research has revealed dangerously high cadmium levels in beef, veal,

0028-0836/81/060436-02$01.00

fish, cereals and potatoes, as well as in some types of mushrooms and mussels. Both Danish and German reasearchers thought that the cadmium in foodstuffs came mainly from sewerage works through the use of the end product as fertilizer.

The European Commission has previously made strenuous efforts to dissuade the Swedes from making a sweeping ban on cadmium's industrial uses. The Swedish government is phasing out "unnecessary" uses of cadmium, but there are not as yet suitable substitutes for cadmium in its many industrial applications.

According to German Home Minister Gerhardt Baum, industry will either have to restrict its use or be faced with a legally enforced ban. The Commission's proposals, if they are finally adopted, will become national law in any case, but as they stand they only serve to protect the aquatic environment, falling short of the restrictions likely in Sweden and Germany.

Two other draft directives on discharge of chemicals into the aquatic environment are already with the council, one on mercury and the other on the "drins" (dieldrin, aldrin and endrin). Other directives are to follow but so far none have been adopted, because member states have failed to agree on limiting values and suitable environmental quality objectives.

The new directive on cadmiun will add to the confusion, particularly as it is in two stages, to come into effect in 1983 and 1986, in recognition that the technology for removing cadmiun from effluents has still to be developed. One major problem yet to be solved is how to remove cadmium from the wastes created during the manufacture of phosphoric acids from phosphate ore. So until a viable technique is developed, the restrictions will not apply to this industry.

Jasper Becker

Soviet psychology

IQ rehabilitated?

An article in the November/December issue of the Soviet Academy of Sciences publication *Psychological Journal* could pave the way for a controversial about-face by Soviet authorities on the issue of intelligence testing.

Stalin made IQ tests illegal in the Soviet Union in 1936. Before that, over-enthusiastic and under-qualified Russian educational psychologists, working with Western tests not standardized on the local population, had designated massive numbers of rural schoolchildren educationally subnormal — to the dismay and outrage of their parents and teachers.

The article in *Psychological Journal*, under the title "The evolution of cognitive processes and abilities", is likely to be especially influential because its author is Friedhard Klix, president of the International Union of Psychological Sciences and director of the psychology section of the Humboldt University, East Berlin. The text formed the presidential address to the opening of the Eastern bloc's 22nd International Psychology Conference in Leipzig in July 1980.

The article stresses the need to treat human intelligence as a subject for study in exactly the same way as other biological phenomena. This is in itself a heresy, since Marxist-Leninist theoreticians have always stressed the need to avoid what they called "reductionism", particularly in relation to aspects of behaviour such as intelligence.

Klix places human intelligence firmly in a biological context and subject therefore to the laws of evolutionary genetics; thereby he accepts an indisputable element of heritability. He refers scathingly to the choice (carefully backdated in his example) that has always faced man in attempting to predict and control his own environment — that is, between adherence to a "cult" and the (scientific) alternative of methodical observation. The unspoken inference is that scientists must be allowed to work unhampered by the dogma of political theory.

Klix later refers (under cover of a paragraph concerned with Sumerian mathematical innovation) to recent experiments in his own university department aimed at pinpointing differences in performance between mathematically gifted children and "normals". The Soviet "ukaz" against objective testing, and in particular the prevention of the development of quantitative norms agains which children with learning difficulties or exceptional gifts can be measured, is well known to Klix's audience.

Western observers feel that the demands of the Soviet-styled "scientific and technological revolution" may show that this prohibition works against rather than in favour of the very children whom it was intended to protect — selection for special education in one of the schools for gifted children at present depends on performance in one of the so-called "Olympiads". These events are held all over the country, but regularly return an embarrassing over-representation of urban males. The use of tests capable of revealing intelligence potential, uncrystallized by access to sophisticated teaching methods or the enriched cultural opportunities of the city, might restore balance to the selection process.

Klix ends by pointing out that the task before psychology is to achieve for cognitive processes what biochemists have managed in their microanalysis of other biological phenomena. Soviet psychological journals-watchers will be agog to see whether the debate on the "microanalysis" of human intelligence along the lines suggested by Klix is continued within the profession. At least the East Germans — who have been quietly using their own intelligence tests for selection for many years — have provided the Russians with an acceptable (that is, foreign) example on which to base the debate.

Elizabeth Roberts

Orlov hunger strike

Dr Yurii Orlov, the Russian physicist, who is now serving a 7-year prison-camp sentence for his activities as chairman of the Moscow "Helsinki Monitoring Group", has gone on a hunger strike to coincide with the reconvening of the Madrid review conference last week. According to his wife, Irina, Dr Orlov made a similar protest last November, when the conference opened, although, during that month, he served two fifteen-day terms in the punishment cell after an alleged argument with a camp officer.

Yurii Orlov — now a fast

The Madrid Conference on Security and Cooperation in Europe opened last week with the prospect of two months of acrimonious debate (see *Nature* 29 January, p.343).

Dr Orlov's protest, said Irina, is an appeal to all governments represented at Madrid to grant amnesty to political prisoners and to reduce secrecy in the flow in all fields of information — social, economic and military.

So far, however, the Soviet Union seems unwilling to relax its restrictions on information flow. During the Madrid recess, two scientists, Aleksandr Lavut, a Moscow geologist, and Mark Niklus, an Estonian biologist, received sentences of 3 and 15 years respectively for publicizing abroad what they see as breaches of the Helsinki Accords on human rights. "Secrecy", too, is a frequent Soviet excuse for not allowing would-be emigrants to leave the country. The latest to be denied an exit visa on the grounds that he had had access to classified information, is Dr David Goldfarb, who until he resigned his post last year when filing his emigration application had been head of the Laboratory of Molecular Genetics of Bacteria and Bacteriophages of the Soviet Academy of Sciences.

0028-0836/81/060437-01$01.00

CORRESPONDENCE

Laser war games

Sir — Zoological and anthropological evidence suggests that old conflict systems never die — they are ritualized. Antagonistic relationships are maintained in principle, but rendered progressively less dangerous as destructive behaviour gives way to display.

If increasing proportions of United States and Soviet military budgets are to be devoted to "laser weapons" (*Nature* 15 January, pages 109–111), this trend should be welcomed as a real sign of international relaxation. Strikingly coloured laser beams of spectacular power can be directed harmlessly into space at regular intervals, with none of the unfortunate ambiguity of intention and inelegance of effect which attend the launching of nuclear missiles. In due course, as Capitalist and Communist take their honoured places in history with Royalist and Roundhead, the "Firing of the Lasers" will no doubt become a public spectacle in the same class as the "Changing of the Guard".

Edward Wheeler

Newtown, Australia

Exhibition policy

Sir — The debate launched by Dr Halstead on the present orientation of the Natural History Museum at South Kensington, having centred on philosophical issues, leaves aside what I consider the main point. That is that because space has been made for new exhibits, the quality of which I will not discuss, many of the established collections are no longer on display. This was a shock to me when I revisited this institution last autumn for the first time in many years. As a kid in the 'fifties my parents took me to London every year and the Natural History Museum was my favourite place to visit. Those visits were certainly important in shaping my vocation for systematic biology. I especially enjoyed those Victorian cabinets full of shells and insects that present managers have eliminated. I believe that for a child it is more exciting to open drawers full of bugs than to push buttons to see clades lighting up. I believe a natural history museum should be a display of as representative as possible a sample of the taxonomic diversity of the living world. Didactic exhibits on the causes of this diversity (evolution) or the techniques used by systematicians should complement the collections and if there is no space available for them, such exhibits could just as well go into the Science Museum.

There is certainly less work involved in preparing displays of stuffed or pickled animals than flashy exhibits on modern biology and maybe the increase (Parkinsonian?) in staff specialized in exhibit preparation, as testified by position openings advertised in *Nature*, is more relevant to the disappearance of collections from the public galleries of the British Museum than are ideologies impregnating the staff.

V. Demoulin

Department of Botany,
University of Liège, Belgium

Corrigendum: In the letter "Energetic consensus" from F.H. Malpress, page 335 in *Nature* of 22 January, the units in the third equation should be kcal mol^{-1}, and not kcal.

Sir — How charming that theological and political issues have come back into zoology[1]. Halstead[2] says he is not worried if Marxism creeps into the Natural History Museum, though he does object to "unsubstantiated assertion" while at the same time he objects to different views presented by various sections of the museum — does this not show a due eclecticism on the museum's part?

Halstead's latest contribution[2] is a direct criticism of the museum's Public Services Department. May I mention my own experience with this department, and its head, Dr Roger Miles? Having been asked some time ago to advise on an exhibition concerned with human perception, I later wrote a fairly detailed criticism of the results, in an editorial of the journal *Perception*[3]. The response was not the usual rebuttal, but quite the reverse — Dr Miles immediately wrote inviting me to suggest plans for a complete redesign of this exhibition, and to attend the planning meetings and essentially to write the commentaries. This we have done. Whatever the result when it appears, this is evidence that the museum takes criticism seriously and acts with unusual generosity in seeking the help and advice of people outside its staff. Having also just given the museum's Christmas lectures, and so interacted again with the Public Services Department, I must say that in my opinion they have wholly commendable enthusiasm, and scholarship, as well as willingness to accept and act on well meant criticism.

Possibly scientists and writers who have not had experience designing exhibitions do not realize the difficulties. In a book, one can give graded qualifications, and cite rival views; and books are written for particular readerships whose background knowledge and interests may be assumed. But large public exhibitions must cater for children and expert authorities. They should be authoritative and balanced according to present knowledge, while also stimulating naive questions and deep questions for future research. To achieve this with an absolute minimum of text is possible only within limits, which we would all surely like to extend. The Natural History Museum should not merely be preserved as a museum of itself.

R. L. Gregory

Department of Anatomy,
The Medical School, Bristol, UK

1. Halstead, L. B. *Nature* **288**, 208 (1980).
2. Halstead, L. B. *Nature* **289**, 106–107 (1981).
3. Gregory, R. L. *Perception* **7**, 1 (1978).

Sir — May I, as a voluntary worker in the Education Department of the Natural History Museum, add a word to the current discussion in your pages.

I understand as little of cladistics as some of your correspondents seem to, but I wish to put in an ardent plea against the steady destruction of the museum by the Public Services Department. The present policy of tearing up by the roots so many of the excellent, clear, logical and didactically ideal displays and replacing them with gimmicky exhibitions that certainly amuse, but do not instruct, is painful to observe. All is now flashing lights, obscure working-models (all too frequently out of order from over-use by the under-tens), complicated texts that appear on screens only to disappear when you take your finger off the button before having read or digested them and strange puzzles that nobody has yet understood. These new, trendy, expensive displays, in which, as D. T. Donovan so rightly points out (*Nature* 1/8 January, p.105), original specimens hardly figure or are replaced by colour-photographs or plastic models, reach their tasteless peak in the disco/fun-fair atmosphere of the Human Biology Hall and provide mainly pseudo-information from which little is learnt and even less retained.

I must except from this criticism the excellent new ecology exhibit. It is a pity that it has been relegated to a long narrow hall and disposed in zig-zag fashion so that visitors are constantly colliding as they slalom their way from side to side, missing many of the showcases in the process.

The loss of the remarkable comparative anatomy section as a teaching-aid on adaptation is a tragedy . . . or is its disappearance really a confirmation of more sinister motives?

As for the scattering of the dinosaurs, a logical teaching tour has become a magical mystery tour, with specimens appearing and vanishing from week to week. Poor *Tyrannosaurus rex* remains in splendid isolation in the old hall; if you can find him. Our splendid *Diplodocus* is completely dwarfed by the enormous proportions of the great central hall, yet he cannot even stretch out his tail as in the old gallery, and his head, inserted with difficulty under the bridge, all but protrudes out of the front entrance, where it can only be studied by obstructing and annoying the incoming visitors. This break-up of the fossil collections, for which the right wing of the museum was specifically designed, is a major disaster and if they really mean to dismantle the fossil mammal gallery, I for one will give up and go, much as I love the work.
(Name and address withheld — ED. Nature)

Soviet rumbles

Sir — Ms Vera Rich has given an inaccurate account of the status of plate tectonics in the Soviet Union (see *Nature* 14 August 1980, p.652). Professor V. V. Beloussov (who is a Corresponding Member of the Academy, not a full Academician, and my close personal friend) has no control over what his fellow members of the Academy write or state about plate tectonics. Academician Peyve, also a close personal friend and director of the Institute of Geology of the Academy of Sciences, has published pro-plate tectonics articles with more than 30 of his colleagues since 1966. There are no restrictions whatsoever within the Academy against publishing papers on plate tectonics, and more than 3,000 pro-plate tectonics articles a year are published in the Soviet Union by members, not only of the Academy of Sciences and its branches, but also of regional institutes and academies.

I am an American citizen, not a Soviet citizen, and am fairly well known for my views against plate tectonics. However, I wish to see the scientists of the Soviet Union given their due. Some of the finest pro-plate tectonics literature in the world is being published in the Soviet Union.

A. A. Meyerhoff

Tulsa, Oklahoma

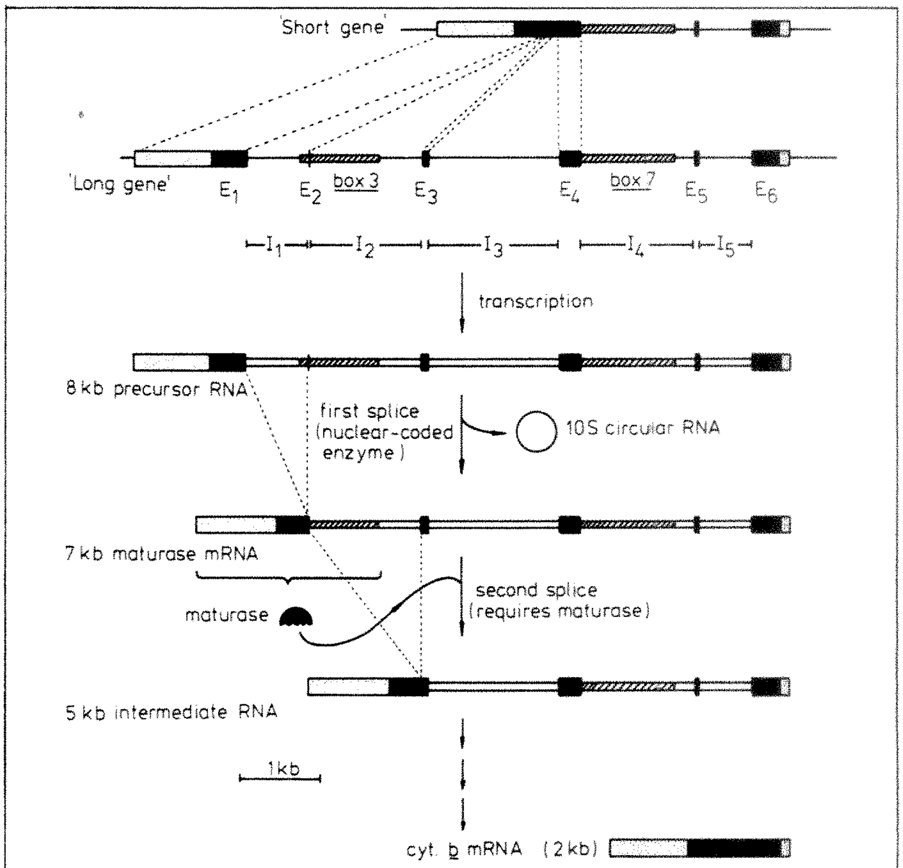
One gene's intron is another gene's exon

from Piet Borst and Leslie A. Grivell

In 1977 Slonimski and co-workers[1] predicted from an impressive genetic analysis that the gene for apocytochrome *b* in yeast mitochondria would be split and that some of the introns would encode separate functions required for the synthesis of *b* (see review in ref. 2). The split nature of the gene was rapidly confirmed by electron microscopy of DNA–RNA hybrids between this gene and its major transcript, an 18S RNA with mRNA activity for *b* (refs 3,4). The nature of the intron-encoded functions remained elusive, however. Initially, Slonimski pushed the idea that the introns code for 'guide' RNAs, that is, RNAs complementary to exon–intron boundaries in precursor RNA and required for removal of intron sequences[2,5]. More recently, however, several groups[6–12] inferred from indirect evidence that some of the introns in the *b* gene code for proteins required for RNA splicing. Although the splicing activity of these proteins remains to be demonstrated directly, a recent paper from Slonimski's laboratory[13] goes a long way towards proving that their synthesis is essential for the proper processing of *b* gene transcripts.

Our present knowledge of the *b* gene in yeast mitochondria is depicted in the figure. There are two versions of the gene in laboratory yeast strains[4,14,15]. The 'long' gene contains six exons (E1–6), five introns (I1–5) and two long open reading frames in introns (hatched) which are continuous with the preceding exon. The 'short' gene lacks the first three introns. An early step in the processing of the primary transcript of the long gene is the removal of I1, giving rise to a small circular RNA (10S, refs 5,11,16). Although the exact site of fusion is now known, the effect of this step is to link exon E1 with the tiny (14-base pair) exon E2 (ref. 13) and the open reading frame fused to it. This allows the synthesis of a fusion protein of molecular weight 42,000, called the *box3* RNA maturase by Lazowska *et al.*[13]. This protein helps to catalyse the excision of the I2 intron and therefore destroys the mRNA which directed its synthesis. Mutations that interfere with maturase function prevent this destruction and lead to overproduction of defective maturase.

The presence of a long open reading frame in the I2 intron was first inferred from the observation that *box3* mutants produce new proteins that look larger than *b* and react with antibodies against *b* (refs 6,7). Since the splicing of precursor RNAs is blocked in these mutants beyond the excision of the 10S RNA circles (see figure), these proteins must arise by readthrough of ribosomes from E1 into a larger open reading frame in I2. This has now been confirmed directly by Lazowska *et al*[13] through sequence analysis of the intron. They go on to show that splicing deficiency in three *box3* mutants is in each case correlated with sequence alterations in the I2 reading frame, both missense and nonsense mutations being effective in preventing splicing. Some of these mutants can be suppressed by paromomycin[17], presumably because this drug induces



The structure of the 'long' and 'short' versions of the gene for apocytochrome *b* and of the mRNAs transcribed from the 'long' gene. Black boxes correspond to the exons E1–E6 which specify the amino acid sequence of *b*; dotted blocks correspond to the non-translated regions of *b* mRNA; hatched blocks represent long open reading frames in the introns of the *b* gene. Whether the 8-kilobase precursor RNA is the primary transcript or not is uncertain. See text for further explanation. Redrawn from refs 4, 13 and 14. Kb, Kilobases or kilobase pairs.

Piet Borst and Leslie A. Grivell are in the Section for Medical Enzymology and Molecular Biology of the Laboratory of Biochemistry, University of Amsterdam.

0028-0836/81/060439-02$01.00

sloppy reading of mRNAs by mitochondrial ribosomes. Since maturase is thought to act catalytically in promoting splicing, a minute amount of misreading should be sufficient to overcome the *box3* mutation.

These results prove that translation of the open reading frame in I2 is required for the proper splicing of precursor RNAs transcribed from the long *b* gene. Two further observations make it clear that maturase is a chimaeric protein, containing the first 143 amino acid residues of cytochrome *b* fused to the I2 reading frame. First, the defective maturase overproduced in *box3* mutants reacts with antibodies against *b* (refs 6,7). Second, chain-terminating mutations in the E1 exon abolish synthesis of maturase[10,13]. The E1 part of the maturase displays high hydrophobicity, typical of integral membrane proteins like *b*, but may not contribute much to its function since missense mutations in E1 that result in synthesis of defective *b* do not interfere with maturase function[10,13].

Whether maturase is a splicing enzyme or only a protein that influences the specificity of other splicing enzymes is not known. In wild-type yeast the protein is present in such low concentrations that it cannot be detected by pulse-labelling (ref. 13 and J. Kreike, personal communication). As pointed out by Lazowska *et al.*[13], it should be possible to find mutants that over-produce wild-type maturase and a full characterization of this protein should not be long in coming. From the DNA sequence it is clear, however, that the *box3* part of the maturase is rather unusual. It is very basic and hydrophobic, and 170 of the 280 amino acid residues are specified by pure AU-codons (Phe, Leu, Ile, Tyr, Asn, Lys). Yeast mtDNA has a mole per cent G + C of 18 and previous work has shown that there must be a strong selective pressure on this genome to maximize A + T content. In genes coding for respiratory chain components there is a strong preference for A or U in the third position[14]; the rRNAs are with 23 per cent G + C the most AU-rich known and intergenic regions are often more than 90 per cent A + T (see ref. 18). The mitochondrial genes coding for major proteins in yeast still contain 27–31 per cent G + C (see ref. 14). Apparently the requirements for maturase function are lax enough to allow the use of 60 per cent pure AU-codons, resulting in a gene containing as little as 18 per cent G + C.

A novel feature in the scheme depicted in the lower half of the figure is that maturase controls its own synthesis, a phenomenon which Lazowska *et al.* call 'splicing homeostasis'[13]. It is likely that this phenomenon is not limited to the *box3* maturase system. The I3 intron also contains a long open reading frame[14]; mutations in this frame (*box7* mutants) are blocked in a late splicing step and also produce novel proteins[10,11,13]. The I4 reading frame may therefore provide the C-terminal part of a second maturase. Several open reading frames have also been found in introns in the *oxi3* region of mtDNA which codes for subunit I of cytochrome *c* oxidase[19]. The set of overlapping transcripts originating from this region is even more complex than the transcripts of the *b* gene[11,20]. This region may therefore give rise to several additional maturases. Slonimski[21] has even suggested that splicing homeostasis also occurs outside mitochondria, but this seems doubtful. His proposal requires the existence within nuclei of a protein-synthesizing system capable of reading the opal stop codon UGA as Trp, as mitochondria do (see ref. 22). At present the existence of such a system is hypothetical.

Another remarkable feature of the intricate *box3* maturase system is that it is dispensable. The information in introns I1–3 is present only in 'long' *b* genes and nowhere else in yeast mtDNA (see ref. 3). Hence, strains with short *b* genes make *b* without the intervention of the *box3* maturase. The I1–3 introns belong to a class of 'optional' introns[23] discovered by Sanders *et al.*[15] in yeast mtDNA. Optional introns are also found in the large rRNA gene and in the gene for subunit I of cytochrome *c* oxidase (see ref. 3). Of course, whether an intron is classified as optional depends on finding a yeast strain that lacks it. Since only a limited number of strains have been looked at, all mitochondrial introns may eventually turn out to be optional.

Why should a simple genome like yeast mtDNA with its 70 kilobase pairs have (optional) introns in three genes and none in all the others? It is clear that many of the explanations advanced for the presence of introns in nuclear genes do not apply here (see ref. 23). It seems possible, however, that the introduction of (additional) splicing steps in the processing of precursor RNAs transcribed from key mitochondrial genes could allow a more complete control over mitochondrial gene expression at the RNA processing level. This might be of use in yeast, which can selectively suppress the synthesis of some mitochondrial components, like cytochrome *c* oxidase or *b*, in anaerobiosis or high glucose. This proposal does not explain the existence of an intron in the rRNA gene, however. Another possibility is that the increased length of intron-containing genes allows more effective recombination within these genes. This could be of advantage in rapidly creating new alleles resistant to yeast inhibitors found in nature. Some of these inhibitors affect *b* (antimycin, mucidin), others the large subunit of the mitochondrial ribosome (chloramphenicol, erythromycin).

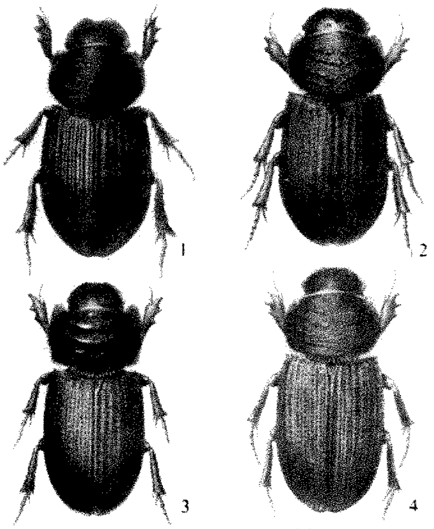
Where do these optional introns come from? In their present form they do not have any of the characteristics of insertion elements; each intron is unique and is only found at a single position in mtDNA. It is possible, however, that optional introns have once arisen from insertion elements and that these elements have lost their mobility in the course of evolution. Bacterial transposons usually contain the information for the synthesis of a transposase, which catalyses the replicative transposition of the transposon (see ref. 24). Transposition involves cleavage exactly at the ends of the transposon, analogous to the precise cuts required for the excision of intron sequences from precursor RNAs. The maturases might therefore have evolved from transposases. A transposon which encodes a maturase which can help to excise the RNA transcript of the transposon from larger precursor RNAs could function as a mobile intron. Such mobile introns would allow the effective introduction of new introns into existing genes. It has been suggested[25] that introns in mitochondrial genes reflect the presence of introns in the ancestral primitive endosymbiotic bacterium from which mitochondria are thought to have evolved. It is also possible, however, that (some) introns were a later acquisition of the mitochondrial genome. This could explain why some introns in the *b* gene have strong homology with introns in the gene for subunit I of oxidase, even though this subunit and *b* have no clear homology themselves[14,19]. Such a late acquisition is also compatible with the fact that the introns in the genes for the large rRNAs of yeast mitochondria and *Chlamydomonas* chloroplasts are in different positions in the gene (see ref. 26). □

1. Slonimski, P.P. *et al.* in *Biochemistry and Genetics of Yeast* (eds Bacila, M., Horecker, B.L. & Stoppani, A.W.N.) 391-401 (Academic, New York, 1978).
2. Dujon, B. *Nature* **282**, 777-778 (1979).
3. Borst, P. & Grivell, L.A. *Cell* **15**, (1978).
4. Grivell, L.A. *et al. ICN-UCLA Symp. molec. cell. Biol.* **15**, 371-384 (1979).
5. Harlbreich, A. *et al. Cell* 321-329 (1980).
6. Kreike, J. *et al. Eur. J. Biochem.* **101**, 607-617 (1979).
7. Solioz, M. & Schatz, G. *J. biol. Chem.* **254**, 9331-9334 (1979).
8. Jacq, C., Lazowska, J. & Slonimski, P.P. *C. r. hebd. Séanc. Acad. Sci., Paris* **290**, 89-92 (1980).
9. Church, G. & Gilbert, W. in *Mobilization and Reassembly of Genetic Information* (eds Joseph, D.R., Schultz, J., Scott, W.A. & Werner, R.) (in the press).
10. Alexander, N.J. *et al. Cell* **20**, 199-206 (1980).
11. Van Ommen, G.J.B. *et al. Cell* **20**, 173-183 (1980).
12. Haid *et al. Curr. Genet.* **1**, 155-161 (1980).
13. Lazowska, J., Jacq, C. & Slonimski, P.P. *Cell* **22**, 333-348 (1980).
14. Nobrega, F.G. & Tzagoloff, A. *J. biol. Chem.* **255**, 9821-9837 (1980).
15. Sanders, J.P.M. *et al. Molec. gen. Genet.* **157**, 239-261 (1977).
16. Church, G.M., Slonimski, P.P. & Gilbert, W. *Cell* **18**, 1209-1215 (1979).
17. Dujardin, G. *et al.* in *The Organization and Expression of the Mitochondrial Genome (eds* Kroon, A.M. & Saccone, C.) 157-160) (Elsevier, Amsterdam, 1980.)
18. Bernardi, G. & Bernardi, G. *FEBS Lett.* **115**, 159-162 (1980).
19. Tzagoloff, A. *et al.* in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A.M. & Saccone, C.) 181-190 (Elsevier, Amsterdam, 1980).
20. Grivell, L.A. *et al.* ibid. pages 37-50 (1980).
21. Slonimski, P.P. *C. r. hebd. Séanc. Acad. Sci., Paris* **290**, Sér. D, 331-334 (1980).
22. Maverick Mitochondria *Nature* **287**, 9-10 (1980).
23. Borst, P. in *Proceedings of Mosbach Colloquia* (eds Bücher, Th., Sebald, W. & Weiss, H.) (Springer, Berlin, in the press).
24. Calos, M.P. & Miller, J.H. *Cell* **20**, 579-595 (1980).
25. Doolittle, W.F. *Nature* **272**, 581 (1978).
26. Dujon, B. *Cell* **20**, 185-197 (1980).

Beetles in ballast

from R.A. Crowson

In former times, ships commonly sailed from European ports to North America heavily 'in ballast', filling up with cargo for their return trips. Ballast was generally picked up in the form of sand or shingle from conveniently accessible points on the shore, and discharged in the vicinity of American ports, often in or around the St Lawrence estuary. In the process numerous European terrestrial invertebrates were introduced into a new continent, in which many of them became established. Beetles, probably on account of their exceptional adult longevity and endurance are very well represented among these involuntary colonists, and were particularly discussed by C.H. Lindroth in his book *Faunal Connections between Europe and North America* (1957). One species he did not consider is brought to light in an excellent recent revision* of the American species of *Rhyssemus* and *Trichorhyssemus* (Scarabaeidae). The European *Rhyssemus germanus*, apparently first found by the St Lawrence estuary in 1929, is now recorded at several points around the Great Lakes and seems to be spreading. The same species seems to have become at least temporarily established in several other remote parts of the world, no doubt in the same way. As can be seen from the figure, *R. germanus* is not particularly similar to any of the native American species, suggesting that separation of the European and American faunas in this group is fairly old.



1 *Rhyssemus scaber* Haldeman;
2 *Rhyssemus neglectus* Brown;
3 *Rhyssemus germanus* (L.);
4 *Rhyssemus sonatus* LeConte.

R.A. Crowson is in the Department of Zoology, University of Glasgow.

---

*(Gordon and Cartwright *The Western Hemisphere Species of Rhyssemus* and *Trichorhyssemus*, Smithsonian Contributions to Zoology **317**, Smithsonian Instn Press; 1980).

A bright nearby supernova

from Roger A. Chevalier

Supernova 1979c was discovered on 19 April 1979 in the galaxy M100 by Johnson (*IAU Circ.* No.3348; 1979). The supernova explosion occurred in a spiral arm of the galaxy implying that it was of Type II — related to the explosions of massive stars. The apparent visual magnitude of the supernova near maximum light was 12, making it the brightest Type II supernova since the supernova 1970g. Observations were made at a variety of wavelengths and have revealed new properties of the Type II supernova emission and enabled a new distance estimate to be made for the Virgo cluster of galaxies.

An initial series of coordinated observations at optical, ultraviolet, X-ray and radio wavelengths was undertaken by an international group of collaborators in the first six weeks after the discovery (Panagia *et al. Mon. Not. R. astr. Soc.* **192,** 861; 1980). The observational facilities included the 5-metre telescope on Mt Palomar, the IUE (International Ultraviolet Explorer) satellite and the VLA (Very Large Array) radio interferometer. The optical properties showed the usual characteristics of a Type II supernova: most of the light was in continuum radiation that could be approximately fitted by a blackbody spectrum with a temperature close to 10,000K. Superposed on the continuum were broad lines of hydrogen and other elements in low stages of ionization. The lines developed P Cygni profiles of emission with blueshifted absorption. If the lines are formed close to the supernova photosphere then the velocity of gas at the photosphere can be deduced from the blueshift of the absorption minimum to be about 8,000 km s^{-1}.

The IUE observations provided the first UV observations of a Type II supernova. The light was dominated by the continuum radiation, although there was excess emission over the blackbody spectrum at low wavelengths. This may be emission from gas heated in the fast shock wave moving into the material surrounding the exploding star (Branch *et al. Astrophys. J.* in the press). Line features were also observed. Of particular interest were symmetric lines of highly ionized atoms (such as N v, Si iv and C iv) with maximum line widths corresponding to velocities of 4,000 km s^{-1}. The lack of occulting effects in the line profiles implies that the emission was at least two photospheric radii from the centre of the star. Panagia *et al.* suggested that the emission was from a circumstellar shell around the presupernova star which had been accelerated by the supernova radiation.

In the months after maximum, the visual light declined with a time scale of about 20 days. Spectra taken in November 1979 and thereafter showed that the light was dominated by a broad Hα line (Branch *et al.;* Kirshner and Chevalier, in preparation). An energy source is required to ionize the gas to produce the Hα line and two possibilities are γ rays from the radioactive decay of isotopes synthesized in the supernova explosion and high-energy emission from a central pulsar. At later times, the IR emission dominated the optical emission. In January 1980, Merrill (*IAU Circ.* No.3444; 1980) found an IR flux about 60 times the flux in the Hα line. He suggested that dust formation occurred in the supernova material, but the mechanism that maintains the ionization in the expanding envelope probably also maintains the gas temperature above the condensation temperature for dust formation. Another possible source of the IR radiation is emission by dust in a circumstellar shell that is heated by the supernova radiation (Bode and Evans *Mon. Not. R. astr. Soc.* **193,** 21; 1980). This model requires an absorption optical depth through the circumstellar shell that is close to the upper limit on the absorption set by observations of the supernova near maximum light, but the properties of the shell are in accord with the known properties of mass loss from red supergiant stars, which are thought to be the progenitors of Type II supernovae.

Radio observations with the VLA in April 1979 did not yield a detection, but gave an upper limit of 0.3 mJy (1 mJy = 10^{-26} erg cm^{-2} s^{-1} Hz^{-1}). However, the observations were repeated in April 1980 at a wavelength of 6 cm and the supernova was clearly detected at a level 5 mJy (Weiler and Sramek *IAU Circ.* No.3485; 1980). The minimum brightness temperature of the emitting region was several times 10^9K if the size of the region was comparable with the radius of the supernova envelope. This indicates that the

Roger A. Chevalier is at the Leander McCormick Observatory of the University of Virginia.

0028-0836/81/060441-02$01.00

emission was nonthermal. It is likely that Type II supernovae do leave pulsar remnants, which are strong nonthermal emitters. However, the observations of the Hα line at the same time as the radio observations indicate that there was sufficient ionized gas to give a large optical depth to free-free absorption. The emission was more likely to be associated with the outer parts of the supernova envelope, but, assuming it was incoherent synchrotron emission, the source of the relativistic electrons and of the magnetic field is not clear.

Besides giving information on the exploding star, observations of the supernova photosphere lead to another important parameter: the distance to the supernova. The technique is to estimate the angular diameter of the supernova photosphere at more than one time from the effective temperature and the total flux from the supernova. The rate of change of the angular diameter, when compared with the actual velocity deduced from the photospheric absorption lines, yields the distance to the supernova. Applying this method Branch *et al.* obtain a distance of 23±5 Mpc. The distance to M100 is especially important because it appears to be a member of the Virgo Cluster, one of the nearest large clusters of galaxies. Distance estimates using more conventional techniques, which depend on nearby distance calibrators, yield distances in the range 12 to 22 Mpc. Distance determinations to supernovae, which are independent of nearby calibrators, should play a fundamental role in determining the distance scale in the universe. □

Negative cooperativity at the insulin receptor

from Alexander Levitzki

In the past five years P. De Meyts and his colleagues have accumulated an impressive amount of experimental evidence suggesting that negative cooperativity is seen in the binding of insulin to its receptor, that is, as the proportion of cell-surface receptors occupied by insulin increases there is a progressive decrease in the affinity of individual receptors for insulin. In the past two years, however, several reports have questioned this concept and, because the behaviour of insulin may be representative of other peptides and proteins binding to cell-surface receptors, it is of some interest to assess its current status.

De Meyts *et al.* have obtained evidence in favour of negative cooperativity both from an equilibrium binding approach and from measuring the rate of dissociation of ^{125}I-insulin as a function of increasing receptor occupancy. Furthermore, in a study of twenty-four different insulin analogues they showed that while some insulins bind with negative cooperativity, others do not reveal negative cooperativity in either binding or in their kinetics of dissociation in the same experimental system[1]. The recent objections to invoking negative cooperativity as the explanation of the data accumulated by De Meyts *et al.* have been of three main kinds.

First, several observations seem to suggest that the enhanced ^{125}I-insulin dissociation at increasing occupancy is not an unambiguous proof of the negatively cooperative model. For example, ^{125}I-TSH (thyroid-stimulating hormone) binds in a non-cooperative manner to cultured thyroid cells, giving linear Scatchard plots, but its rate of dissociation from the receptor is enhanced by increasing concentrations of the unlabelled hormone[2]. Similarly, the enhancement at 10 °C of the dissociation of ^{3}H-dihydroalprenolol from frog β-receptors by native alprenolol occurs at a range of ligand concentration that yields non-cooperative binding (Hill coefficient of 1.0)[3]. In addition, recent studies by Pollet *et al.*[4] reveal no negative cooperativity in the same frog erythrocyte β-receptor system studied by Limbrid *et al.* An enhanced rate of dissociation has also been observed for biologically inert substances. For example, insulin enhances the rate of dissociation of ^{125}I-insulin from talc[5] and similarly, the rate of dissociation of ^{125}I-abrin from Sephadex particles is enhanced by unlabelled lectin or lactose[6]. In addition, difficulty arises from the claim of Pollet *et al.*[7] that no enhancement of ^{125}I-insulin dissociation is obtained at increasing receptor occupancy in the same IM-9 lymphocytes as those studied by De Meyts *et al.* The latter findings are in direct conflict with those reported by De Meyts and his colleagues[8] which were recently confirmed, on the same cell line, by Sonne and Gliemann[9]. The main conclusion from these different studies is that one must recognize that the existence of an enhanced dissociation of a ligand from its receptor in the presence of excess ligand, in cases where negative cooperativity clearly does not occur, eliminates this method as a definitive tool. It has been recently suggested how it might be possible to obtain a definitive interpretation by means of a new type of equilibrium measurement[10].

A second objection is raised by a recent study which claims that the existence of high- and low-affinity sites represents heterogeneity and not negative cooperativity[11]. The authors demonstrate that the binding isotherm of ^{125}I-insulin to human red cells is fitted well by a model which assumes two classes of sites: each cell reveals one high-affinity site ($K_d \sim 71$ pM). The authors further show that the high-affinity site can be completely blocked by the binding of concanavalin A, leaving the low-affinity sites intact. This finding is not compatible with the negative cooperativity model of De Meyts *et al.* One must remember, however, that a sample of human erythrocytes is always heterogeneous, certainly in their age. It would not be surprising if one were to find that the high-affinity sites and the low-affinity sites do not reside on the same erythrocyte, in which case the entire argument of Herzberg *et al.* is invalid. It is also difficult to use an experimental system that exhibits ony one high-affinity site per cell and in which it is not clear whether the insulin receptors are functional.

The final complication was pointed out recently by Donner[12]. ^{125}I-insulin is degraded rapidly by intact hepatocytes into ^{125}I-insulin fragments. These fragments possess a spectrum of affinities that is manifested in curvilinear Scatchard plot for ^{125}I-insulin binding to the intact hepatocytes. When the authors correct for the true intact ^{125}I-insulin, the binding pattern is converted into a non-cooperative one, represented by a linear Scatchard plot. Thus, when one does not correct for ^{125}I-insulin, one observes about 30,000 high-affinity sites per cell ($K_d \sim 1.0$ nM) and about 600,000 sites per cell ($K_d \sim 60$ nM). When such a correction is made, only 36,000 sites per cell ($K_d \sim 0.4$ nM) are observed. These results also fit with the common observation that the biochemical response of these cells to insulin is of a non-cooperative Michaelian type. The authors, in fact, confirm earlier reports[13] in which it was reported that ^{125}I-insulin is degraded to fragments by intact hepatocytes.

Although Donner's findings are relevant to intact hepatocyte cells which are obviously loaded with degradative

---

Alexander Levitzki is Professor of Biochemistry at the Hebrew University of Jerusalem.

1. De Meyts *et al. Nature* **273**, 504 (1978).
2. Verrier *et al. Eur. J. Biochem.* **42**, 355 (1974).
3. Linbrid *et al. Biochem. biophys. Res. Commun.* **64**, 1160 (1975).
4. Pollet *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4430 (1980).
5. Cuatrecasas & Hollenberg *Biochem. biophys. Res. Commun.* **62**, 31 (1975).
6. Sandrig *et al. Eur. J. Biochem.* **88**, 307(1979).
7. Pollet *et al. J. biol. Chem.* **252**, 5828 (1979).
8. De Meyts *et al. Biochem. biophys. Res. Commun.* **55**, (1976); *J. biol. Chem.* **251**, 1877 (1976).
9. Sonne & Gliemann *J. biol. Chem.* **255**, 7449 (1980).
10. Levitski *et al. Eur. J. Biochem.* **103**, 449 (1979); *Proc. natn. Acad. Sci. U.S.A.* **77**, 5055 (1980).
11. Herzberg *et al. Nature 286*, 279 (1980).
12. Donner *Proc. natn. Acad. Sci. U.S.A.* **77**, 3176 (1980).
13. Freyetet *et al. J. biol. Chem.* **247**, 3953 (1972).

0028-0836/81/060442-02$01.00

enzymes, his conclusion cannot be extrapolated to the purified membrane systems studied by De Meyts and his colleagues, and most probably does not apply to other cell systems. Indeed, only recently Sonne and Gliemann[9] have shown that ^{125}I-insulin is not degraded at all by the IM-9 lymphocytes, even at 37 °C. This cell line exhibits strong negative cooperativity towards ^{125}I-insulin, as judged both by binding studies and by accelerated dissociation of ^{125}I-insulin[1,9] at all temperatures including physiological ones.

To summarize, the most important point raised by the critics of the negative cooperativity model for insulin receptors is the demonstration that the accelerated rate of dissociation of a ligand from its receptor at increasing occupancy is not a definitive proof for negative cooperativity. Thus, in order to demonstrate negative cooperativity (or its absence) one must resort to other methods as well. For insulin itself, it seems that the experimental evidence for negative cooperativity in quite a number of systems is still very convincing. □

The Trigger experiment: simulating auroral processes

from Daniel W. Swift and Hans C. Stenbaek-Nielsen

TO INVESTIGATE the possibility of triggering auroral processes, a rocket experiment, appropriately code named 'Trigger', was designed and carried out by a group of Swedish and United States auroral researchers. In Trigger the electrical properties of the ionosphere were artificially perturbed by an explosive release of caesium metal and the effects were observed by a wide range of both rocket-borne and ground-based instruments. The experiment took place in February 1977 with launch from Esrange near Kiruna, Sweden, and indeed, the results have recently appeared in the *Journal of Geophysical Research* (10 October, 1980).

Traditionally, the ionosphere has been considered an entirely passive medium and the aurora to be little more than the ionospheric imprint of processes taking place out in the magnetosphere. But the ionosphere is electrically coupled to the magnetosphere with large currents flowing between the two regions and recent data indicate that these two currents are important for auroral processes. Thus it would seem reasonable to expect that events taking place in the ionosphere may affect magnetospheric processes. Evidence for a less passive ionosphere has also been found in observations on some chemical releases, although these experiments were never intended to produce such effects; an example is pulsating auroras apparently stimulated by a barium release (Deehr and Romick *Nature* **267**, 135; 1977).

While we have long known that the optical emissions forming the aurora are for the most part due to energetic electrons precipitating into the upper atmosphere, it is only within recent years that we have begun to understand the mechanisms which are responsible for this precipitation and the role played by the ionosphere.

It appears that two general processes are responsible for the precipitation of auroral electrons. The first process is the result of leakage of energetic electrons from the Earth's radiation belts into the atmosphere. This leakage is greatly enhanced by very low-frequency radio emissions and electrostatic waves which can be generated spontaneously within the radiation belts or stimulated from the outside. Processes of this type were among the earliest suggested and are to a large extent responsible for the diffuse aurora.

The other category of processes is associated with the discrete aurora, for example, the spectacular arcs and bands often observed in the late evening and along the poleward edge of the auroral oval. The main characteristic is potential differences set up along the magnetic field lines in which electrons are accelerated in much the same way as in the electron gun of a television tube. These potential structures are located in the near-Earth magnetosphere and their existence has only recently been verified experimentally.

Although the Trigger experiment did not actually create an artificial aurora, measurements of energetic electron fluxes in the neighbourhood of the plasma release gave indications of both types of auroral precipitation processes.

The caesium plasma was created by a 12-kg mixture of TNT-AlO-CsNO$_3$ which was explosively released at an altitude of 164 km. Caesium is a metal that easily ionizes to form a plasma and thus the release resulted in the rapid motion of an electrically conducting plasma cloud across the Earth's magnetic field. The motion of a conducter across the magnetic field lines constitutes an electrical dynamo which generates electrical voltages capable of accelerating electrons. In addition, the Trigger experiment creates a number of waves which may travel out into the magnetosphere and it leaves in its wake an enhanced level of ionization that can also affect plasma processes which occur on magnetic field lines passing through the plasma cloud.

An instrumented payload was sent up with the payload containing the explosive charge, but separated from the charge before the burst, so that when the explosion occurred, it was separated by 1.5 km and on a magnetic field displaced by about 0.5 km from the instrumented payload. A large peak in the electric field

Photograph taken from the optical observing site at Andoya Rocket Range, northern Norway, of a natural aurora showing auroral arcs (with a little local cloud in front of them), which stretched from the eastern horizon to the western horizon (behind the mountains). (Taken by H.L. Collin of the Geophysical Institute.)

The authors are in the Geophysical Institute at the University of Alaska.

0028-0836/81/060443-02$01.00

was seen about the time the blast wave should have crossed the magnetic field line intersecting the payload. Nearly coincident with this, electron fluxes of 2-kV were seen moving up and down the field lines, but the electrons did not actually seem to reach the atmosphere. The currents generated by the electric field apparently set up reflecting electric shocks both above and below the burst point, and the electrons were probably trapped between them.

The discrete auroras are believed to be caused by electrons accelerated through similar shocks, or double layers, also set up by currents flowing parallel to magnetic field lines; but although the current-driven shocks play roles in both the natural aurora and the Trigger experiment, their roles are quite different.

The instrument package also observed electron fluxes with energies greater than 40 kiloelectron volts (keV). There was a prompt burst followed a second later by sustained fluxes of electrons. Fluxes greater than 40 keV are believed to be responsible for the patchy and pulsating aurora seen in the early morning hours in the diffuse aurora (see review by A.D. Johnstone *Nature* **119**, 274; 1978). The natural precipitation is also accompanied by electromagnetic wave emissions called chorus. The chorus emission is in the audio frequency range, 0.5 to 2 kHz, and when amplified sounds like the chirping of birds. The precipitation can be enhanced by increasing the ambient ionization because the ambient ionization makes the trapped electrons more unstable to precipitation caused by the electromagnetic emissions.

It is tempting to attribute the delayed precipitation of electrons of energy greater than 40 keV in Trigger to the enhanced ionization density left by the caesium in the wake of the Trigger blast wave. The delayed fluxes were accompanied by wave emissions, but the emissions were electrostatic, rather than electromagnetic, and so would not probably be able to cause precipitation of trapped fluxes. Moreover, the Trigger fluxes were primarily field aligned, which is again inconsistent with a wave-particle diffusion process. The experimenters attribute most of the enhancement to noise contamination of the magnetic field-aligned detector. However, an X-ray detector suspended from a parachute that was deployed by a second rocket did observe 0.13-Hz pulsations in X-ray fluxes some two minutes after the event. This frequency is comparable with pulsation frequencies observed in natural aurora.

The prompt energetic electron fluxes are believed to be real, and they are accompanied by an electromagnetic pulse. However, as the electron fluxes are also primarily field aligned, the acceleration mechanism would have to be a stochastic process, since the electric potentials generated by the explosion are considerably less than the observed electron energies.

The Trigger experiment initiated a number of interesting plasma processes, some of which are similar to those responsible for the aurora. However, the natural aurora and the events observed in Trigger are not similar enough for the Trigger results to be applied in a direct way towards understanding the aurora; but after further theoretical interpretation it may turn out that the experiment has provided a number of useful clues to understanding processes in the natural aurora.

Certainly, the concept of the Trigger experiment is very appealing. Not only does the technique of artifically perturbing the ionosphere or magnetosphere allow the study of a wide range of general plasma processes, it also makes it possible to separate cause and effect, an often serious problem in the interpretation of auroral observations.

The use of chemical releases in space research will increase; groups from many countries are now proposing that such release experiments be conducted from a 'chemical release module' which will be be launched through the US space shuttle programme sometime in the mid-eighties. □

Are membrane proteins introverted?

from N. Michael Green

THE PURPLE MEMBRANE of *Halobacterium halobium* has become a test bed for sophisticated new techniques. These, in turn, have given exciting new information about the structure and function of the major protein of the purple membrane, bacterial rhodopsin. The most recent technique, neutron diffraction, shows that rhodopsin is an "inside-out" protein. Six years ago a combination of novel electron microscopic techniques with X-ray diffraction gave a picture of the structure at 7 Å resolution (Henderson and Unwin *Nature* **257**, 28; 1975). Then last year, a correlation of the electron density distribution with the amino acid sequence of appropriate helical segments produced a tentative model structure (Engelman *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 2023; 1980). It is this model which now receives strong support from some elegant neutron diffraction experiments (Engelman and Zaccai *Proc. natn. Acad. Sci. U.S.A.* **77**, 5894; 1980). This method, based on the great difference between neutron scattering by protons and deuterons, has also been used to show that there are no detectable aqueous channels in the membrane (Zaccai *J. molec. Biol* **132**, 181; 1979; Rogan and Zaccai *J. molec. Biol.* **145**, 281; 1980.).

Purple membranes were prepared in which all of either the valine or the phenylalanine residues had been replaced by their fully deuterated analogues and the neutron diffraction patterns were compared with that from normal membrane. From the latter a neutron density map was constructed which closely resembled the electron density map. Difference Fourier maps relative to the deuterated preparations were computed and showed that valine and phenylalanine residues were concentrated in different regions. This was to be expected, since these residues have rather different distributions both between the seven transmembrane helices and azimuthally within each helix. The peak density differences for both valine and phenylalanine side chains were off the helix axes, as would be expected, but whereas the valine peaks were displaced towards the exterior of the molecule those of phenylalanine were nearer to the interior, in the region where many of the charged groups had been placed in the tentative structure of Engelman *et al.* This differential distribution was consistent with the relative azimuthal distributions of valine and phenylalanine derived from the sequences of the previously assigned helices.

Comparison of the sequence-based azimuthal distributions with those of the charged residues confirms that valines (17 out of 19) are on the opposite sides of helices from phenylalanines (9 out of 11), and that the buried charged groups go with the phenylalanines. Moreover, the previously assigned locations of helices rich in valine or phenylalanine were consistent with the Fourier difference maps derived from neutron scattering, though the authors refrain from any detailed interpretation of this at the moment. The consistency of these new results with the earlier model is impressive and gives confidence in the validity of the arguments used in support of the previous assignments.

A firm picture is beginning to emerge of an 'inside-out' protein molecule with buried polar and charged groups located predominantly between helices and with most of the interface between lipid methylene groups and the protein being occupied by hydrophobic aliphatic chains. The authors suggest that in addition to a possible role in providing sites for proton translocation, the charged and polar groups would be an important stabilizing element in such a predominantly hydrophobic structure, since in their absence interactions between hydrophobic helices immersed in a hydrophobic solvent are very weak. □

N. Michael Green is in the National Institute for Medical Research, Mill Hill, London.

0028-0836/81/060444-01$01.00

African lithosphere

from Paul Mohr

ONE reason for the temporary demise of continental drift as a respectable scientific hypothesis was that the rigidity of planet Earth appeared too great for such drift to be possible. However, by 1948 the seismologist Beno Gutenberg was able to confidently state that a 'Low Velocity Zone' existed within the Earth's upper mantle. From seismic shear-wave velocity measurements the zone is estimated to extend between depths of about 120 and 200km. Its physical properties are consistent with incipient melting of mantle peridotite rock, and thus the term asthenosphere (zone of weakness) is appropriate. The overlying lithosphere, comprising the crust and topmost mantle, contrasts physically with the asthenosphere in being stronger and more rigid. The lithosphere has become identified with the plates of plate-tectonic theory and these plates are believed to move over the asthenosphere.

The global pattern of motions of the lithospheric plates has now been measured (see ref.1) and the manner of plate generation at sea-floor spreading axes, together with the ensuing thickening of the sub-oceanic lithosphere are well known. Much less well known is the thickness of the lithosphere under continental sectors of plates, and the means by which such lithosphere becomes attenuated where a plate begins to split, as, for example, along the African Rift System.

Early insights into African lithospheric thickness came from surface-wave dispersion studies[2-5] and in 1976 one could say, in summary, that under the undeformed interior of Africa, the top of the asthenosphere lies near 120km depth, but near the margins of the rift system it rises to an average depth of about 80km. Thin though the data were, with the addition of seismic profiling studies of the African crust they provided essential constraints for interpreting the more extensive data derived from other geophysical techniques.

Gravity surveys in eastern Africa have firmly established that lithospheric thinning is associated with the rifting of a continent. Sowerbutts[6] proposed the existence in Kenya of a low-density astenospheric 'cushion' underlying an unthinned lithosphere, but Girdler *et al.*[7] explained the same gravity data solely in terms of lithospheric thinning. More recent work has led to a combination of these two ideas, using a starting thickness for the unattenuated lithosphere of 80-100km[8] or 120km[9]. The asthenospheric 'cushion' beneath the Rift System may have a prism-like cross-section, according both to seismic array[10] and geomagnetic sounding[11] studies, with the lithosphere-asthenosphere boundary initially dipping steeply away from the rift valleys.

A particularly bold contribution to African lithospheric studies has come from combining continent-wide gravity data, based on the 1973 US Defense Mapping Agency's Bouguer-anomaly map of Africa, with observed teleseismic delay-times for Soviet underground nuclear explosions recorded at African seismological stations. Fairhead and Reeves'[12] resulting map of lithospheric thickness, derived from empirical linear relationships among delay-time, station elevation and Bouguer anomaly, reveals a broad zone of thinning coincident with the Rift System, from Ethiopia through Kenya and Tanzania to Angola. To the northwest, across Zaire and the Sahara, the lithosphere purportedly averages 200km thickness (less for the Tibesti-Jebel Marra line), while in southern Africa, southeast of the Rift System, the thickness lies between 175 and 150km. These values are indeed appreciably higher than indicated from the more direct approach of surface-wave dispersion analysis.

An almost identical picture, but with very different absolute values for lithospheric thickness, is most recently provided from a reinterpretation of the US Defense Agency gravity map following on a convincing demonstration of the accuracy of the map[13,14]. Detailed analysis shows that lithospheric thickness associated with the rift valleys and their bordering plateaux is less than 50km, and is less than 70km for a substantial part of southern Africa. These values are significantly less than those derived from seismic-wave studies.

How can such divergent conclusions be reached from essentially the same data base? One obvious reason is that seismological constraints are still not adequate to provide a firm model of density-depth gradients, and thus facilitate a unique interpretation of gravity profiles. Brown and Girdler[14] agree with Fairhead[8] in obtaining a density contrast of 0.05g cm^{-3} at the lithosphere-asthenosphere boundary, but in retaining the fundamental assumption of Girdler *et al.*[7] that the longer-wavelength Bouguer anomalies of Africa are caused entirely by variations in lithospheric thickness. That is to say, there is no contribution to the gravity-field from lateral changes in density below the lithosphere — there is no asthenospheric 'cushion'.

What is the evidence for an asthenospheric 'cushion', that is, for steeper and more profound asthenospheric boundaries than thus far provided by any of the published gravity interpretations? Seismic-array investigations in Kenya indicate that the sub-Rift System asthenosphere remains laterally distinct from 'normal' asthenosphere down to depths of 250km, far below the base of even unattenuated lithosphere[10]. The asthenosphere beneath the Rift System could therefore be hotter, more melted and even less dense than 'normal' asthenosphere under the rest of Africa. Here we approach the crux of the problem.

All agree that volcanism and high heat-flow in Africa are associated with zones of lithospheric thinning. But is this merely the result of passive upwelling of asthenospheric material as the plates begin to separate? This thesis is the rationale for and the explicit conclusion of Brown and Girdler's[14] interpretation of gravity. It is also the conclusion, reached on the entirely different premises of isotope geochemistry of volcanic rocks in Afar, of Barberi *et al.*[15]. Alternatively, are there 'hot spots' or deep mantle plumes impinging on the underside of the African lithosphere, involving vertical mass-transfer that could well derive from the mesosphere at depths of 400km or more[16,17]. This crucial issue is unresolved.

The questions of how thick is the African lithosphere, and the manner of lithospheric thinning as Africa splits, remain to test the best of the next generation of earth-scientists. Further questions arise: How deep and wide is the column of mantle perturbation below the Rift System, and do phase-changes below 400km depth contribute to the gravity field? What is the variation with depth of lateral density gradients within the asthenosphere under eastern Africa? And closer to the surface:

Paul Mohr is Professor of Geology at University College Galway, Ireland.

---

1. Minster, J.B. *et al. Geophys. J.R. astr. Soc.* **36**, 541 (1974).
2. Gumper, F. & Pomeroy, P.W. *Bull. seismol. Soc. Am.* **60**, 651 (1970).
3. Knopoff, L. & Schlue, J.W. *Tectonophysics* **15**, 157 (1972).
4. Long, R.E. *et al. Tectonoplysics* **15**, 165 (1972).
5. Mueller, S. & Bonjer, K.-P. *Tectonophysics* **20**, 283 (1974).
6. Sowerbutts, W.T.C. *Nature* **223**, 143 (1969).
7. Girdler, R.W. *et al. Nature* **224**, 1178 (1969).
8. Fairhead, J.D. *Tectonophysics* **30**, 269 (1976).
9. Wohlenberg, J. in *Afar between Continental and Oceanic Rifting* (eds. Pilger & Rösler) (Schweizerbart, Stuttgart, 1976).
10. Long, R.E. in *Afar between Continental and Oceanic Rifting* (eds. Pilger & Rösler) (Schweizerbart, Stuttgart, 1976).
11. Banks, R.J. & Ottey, P. *Geophys. J.R. astr. Soc.* **36**, 321 (1974).
12. Fairhead, J.D. & Reeves, C.V. *Earth planet. Sci. Lett.* **36**, 63 (1977).
13. Brown, C. *et al. J. geophys. Res.* **85**, 6436 (1980).
14. Brown, C. & Girdler, R.W. *J. geophys. Res.* **85**, 6443 (1980).
15. Barberi, F. *et al. Earth planet. Sci. Lett.* **50**, 247 (1980).
16. Schilling, J.-G. *Nature phys. Sci.* **242**, 2 (1973).
17 Gass, I.G. in *Afar Depression of Ethiopia* (eds. Pilger & Rösler) (Schweizerbart, Stuttgart, 1975).
18. Mueller, S. in *Tectonics and Geophysics of Continental Rifts* (eds. Ramberg & Neumann) (Reidel, Dordrecht, 1978).
19. Thiessen, R. *et al. Geology* **7**, 263 (1979).
20. McKenzie, D.P. *et al. Nature* **288**, 442 (1980).

0028-0836/81/060445-02$01.00

How does plateau crust, thickening as the rift margins are approached, join with the attenuated crust of the rift floors? Does the rift Moho, evolve in nature and depth in the way proposed by Mueller[18]? Finally, is it possible that the detailed gravity signature for Africa expresses sublithospheric mantle convection (as postulated on other grounds by Thiessen *et al.*[19] with horizontal dimensions of about 1,000 km? A rather similar figure has recently been derived from bathymetry-geoid measurements for the Pacific by McKenzie *et al.*[20], hinting that mantle convection beneath oceans and continents is of similar if not identical form. □

Is fibroblast interferon antigenic for man?

from D. A. J. Tyrell

In this issue of *Nature* (p.496) Vallbracht *et al.* report the treatment of a child with interferon β for the malignant disease, neuroblastoma. Interferons are known to act against experimental tumours, possibly by inhibiting cell division and enhancing immune activity, for example by activating natural killer cells. As fibroblast interferon is not usually found in the circulation after intramuscular or subcutaneous injection, possibly because it is bound by tissues, it is given intravenously when a systemic effect is required in clinical studies. In the case reported by Vallbracht *et al.* there was a poor clinical response and while monitoring the pharmacokinetics of the injected interferon it was found that there was no interferon in the blood. An anti-interferon globulin was detected in the circulation, apparently in consequence of the repeated interferon injections. This did not appear in other patients. What does this mean?

From analysis of cloned human genes it is clear that interferons are families of molecules. Apparently there are distinct genes for eight different forms of interferon α (IFN-α or leukocyte interferon), and IFN-β (fibroblast interferon) may also be the product of more than one gene. All interferons do, however, seem to be 'poor' antigens — in the practical sense that it is difficult to make antisera against them.

Nevertheless, antibodies against interferons have been produced by inoculating them into foreign species, and in a recent example a mouse immunized against human IFN-α — exactly which one we do not know — provided a monoclonal antibody-producing hybridoma. Immunoglobulin from the hybridoma has been used to purify interferon for clinical study (Secher and Burke *Nature* **285**, 446; 1980), and may also be used for radioimmunoassay.

One would not expect to be able to produce antibodies by inoculating an interferon into the same species as the one which produced it. In general, immune responses are not mounted against normal constitutuents of cells because of a 'self/non-self' recognition mechanism which distinguishes between normal tissue components and novel and presumably extraneous antigens. However, as a transient response to tissue damage and as part of the 'autoimmune' diseases antibodies can be produced against normal tissue constitutents. This reaction seems particularly likely when the tissue component, for instance thyroglobulin, does not ordinarily enter the circulation and encounter cells of the immune system.

Did the patient treated by Vallbracht *et al.* have some impairment of the self/non-self mechanism and thus react in this way? Is IFN-β normally found only in or near to producing cells, so it is recognized as foreign when it enters the circulation? Was the IFN-β given subtly different from the patient's own IFN-β, perhaps because it was the product of a different gene or because it was modified or denatured during manufacture? It remains to be seen whether anti-interferon antibodies are anything but an occasional nuisance when trying to treat patients with interferon — no such trouble seems to have arisen during the treatment of substantial numbers of patients with IFN-α. However, the single case report stimulates us to ask questions about the biology of interferon which we might otherwise not have thought of asking, and that must be a good thing. □

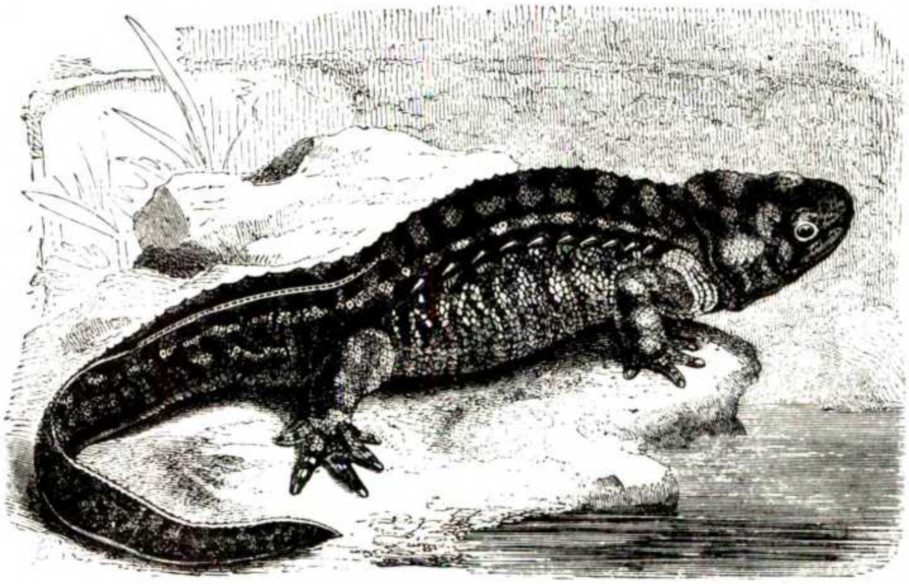
D. A. J. Tyrell is in the Medical Research Council Clinical Research Centre, Harrow, UK.

NATURE

100 years ago

Dr. Lenz, the German traveller who lately accomplished the feat of reaching Timbuctoo from the north, has arrived at Bordeaux, and is expected at Berlin soon to give an account of his explorations.

News from Cairo states that to the north of Memphis, near Saggarah, two pyramids have been discovered which were constructed by kings of the sixth dynasty, and the rooms and passages of which are covered with thousands of inscriptions. The discovery is said to be of the greatest scientific importance.



Pleurodeles waltlii

The habits of the Pleurodeles seem to be more or less like that of our native Tritons. During the procreative season they remain upright in the water; later they leave it and hide themselves in damp places under stones. Like the Water Newts, they posses a sort of cry; when frightened, as on being suddenly seized, they emit a low, short, almost squeaking sound, generally repeated several times. This seemed to come not so much from the throat as to be caused by a rapid expulsion of air through the openings of the nose — in fact, to be a sort of snort.

It would almost seem worth one's while to pay a visit to those Andalusian tanks, and by their semi-limpid sides and under the shelter of their surrounding fig and olive-trees work out the complete history of this interesting little form.

From *Nature* **23**, 3 February, 1881.

0028-0836/81/060446-01$01.00

ARTICLES

Apparent thickness of Saturn's rings

André Brahic & Bruno Sicardy
Observatoire de Paris, Université Paris VII, 92190 Meudon, France

The determination of the physical thickness of Saturn's rings is crucial for a better understanding of the nature, the dynamics and the evolution of a system of colliding ring particles. Ground-based observations with electronographic cameras and CCD during the transit of the Earth through the ring plane in March 1980, reveal a photometric apparent thickness of 1.4 ± 0.3 km. This value is only an upper limit of the local ring thickness, which is not observable from the Earth or from flybys. The observed brightness includes the contribution of the E ring, of large chunks and condensations and of the warping of the disk. Theory predicts a local thickness of the order of only a few mean particles radii.

We know of three dissimilar systems of planetary rings. Saturn has bright, broad rings separated by narrow gaps; the rings of Uranus are dark, narrow and widely spaced while the main jovian ring, closer to its central planet, is intermediate in size and has a three-dimensional halo. These rings are the nearest disk-like systems we have, and should offer valuable insights into the dynamics of more exotic and less accessible flat systems such as spiral galaxies, accretion disks and the primordial Solar System.

Following the discoveries of Uranus' and Jupiter's ring in March 1977 and March 1979, respectively[1,2], Saturn's system is now being intensively observed. Pioneer 11 became the first spacecraft[3] to reach Saturn on 1 September 1979, and the Voyager spacecraft has produced much interesting data, and more discoveries can be expected on its next encounter (27 August 1981). In between these encounters, ground-based observations during the ring plane crossing have revealed new inner satellites[4,5,7], new rings[6,8], and that the surroundings of Saturn are much more 'crowded' than foreseen.

Observations at very low elevations may supply important new information on the shape and the nature of the rings. In particular, when the rings are seen edge-on as a very faint line, the measurement of the brightness of this line may enable us to estimate the physical thickness of the rings.

Photometric determination of the thickness of Saturn's rings

The main problem is the evaluation of the physical thickness z of Saturn's rings, which lies far below the optical resolution of any instrument ($z \sim 1$ km whereas the resolution, from the Earth, is ~6,000 km). Nevertheless, photometric measurements of the rings, even though taken close to a much brighter planet, enable z to be determined.

On the plates we have examined, the Sun and the Earth are on opposite sides of the ring plane, so that the side S we can observe is dark in comparison with the edge E (Fig. 1*a*). The plates we have reduced were taken by Laques *et al.*[9], with the 1.06-m reflector of the Observatoire du Pic-du-Midi, and by Wlérick *et al.*[10], with the 1.93-m reflector of the Observatoire de Haute-Provence. In both cases, the receiver was an electronographic camera built at the Paris Observatory. Table 1 summarizes the reduced observations and provides details about receptors and conditions of observations. A microdensitometry of these plates provides ϕ, the luminous flux per unit length (1 arc s), of the rings (Fig. 1*a*). This flux is compared with that of a satellite which is not overexposed on the plate. Depending on the plate, the reference satellite is Tethys, Enceladus, Hyperion or Mimas. Our photometric measurements agree within an accuracy of 0.15 mag with known data[11].

By using the simplest model of the rings, that is, a plane parallel homogeneous scattering layer with a rectangular cross-section[12] in which light undergoes multiple isotropic scattering, it is then possible to separate ϕ into two parts: ϕ_E, the flux from the edge E, and ϕ_S, from the dark side S. In the ideal case where the rings are just seen edge-on, we obtain $\phi = \phi_E$. Actually this is not observed, so we had to extrapolate $\phi = \phi_E + \phi_S$ from the observed geometries to its minimum value ϕ_E at the edge-on geometry.

For this extrapolation we then need b_E, the brightness of the edge, or the mag/arc s^2 m_E of the edge ($b_E = 10^{-0.4 m_E}$) because $z\,(\text{arc s}) = \phi_E / b_E$.

The tables in ref. 13 provide observed brightness m_v and m_B (mag/arc s^2) of the illuminated side, in visible and blue light respectively at different points of the rings, with a theoretical estimation of τ (perpendicular optical thickness) at these points. The elevation angles B of the Earth and B' of the Sun are then equal to $B \simeq B' \simeq 26°$, and the phase angle $\alpha = 0.094°$. A large optical thickness is derived at the brightest part of B ring ($\tau \simeq 1$),

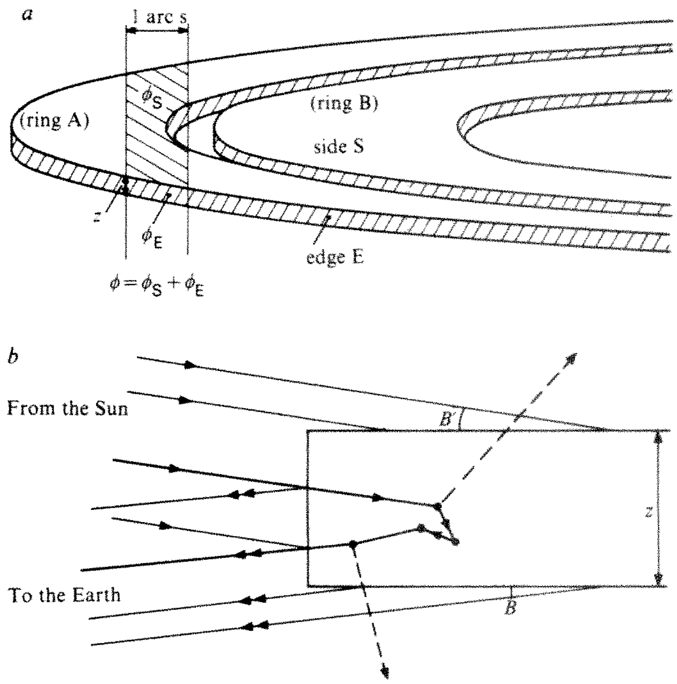


Fig. 1 *a*, Plane-parallel model of Saturn's rings. The thickness has been magnified compared to radial dimensions. *b*, The tilt-angle B is much smaller during the 1980 observations so that the rings appear as a line only.

0028-0836/81/060447—04$01.00

Table 1 Summary of the reduced observations

Identification	Date	μ	μ_0	$B\times B'$ (deg)^2	α (deg)	Colour	Mag/arc s of a ring segment 17 arc s	Mag/arc s of a ring segment 20 arc s	Adopted reference
PbY 136	80 03 12.94	−0.000168	0.00264	−0.00145	0.29	V	15.5 (W)	15.2 (W)	V_0 Tethys = 10.3
PbY 137	80 03 12.95	−0.000178	0.00264	−0.00154	0.29	V	15.6 (E) 15.5 (W)	15.2 (E) 14.7 (W)	V_0 Hyperion = 14.2
PbY 145	80 03 16.01	−0.00220	0.00346	−0.0249	0.34	V	15.7 (W)	14.7 (W)	V_0 Hyperion = 14.2
PbY 152.C	80 03 17.01	−0.00286	0.00372	−0.0349	0.41	V	14.7 (E) 14.9 (W)	14.7 (E) 14.8 (W)	V_0 Hyperion = 14.2
PbY 152.E	80 03 17.01	−0.00286	0.00372	−0.0349	0.41	V	15.6 (E) 14.9 (W)	14.7 (E) 14.4 (W)	V_0 Enceladus = 11.8
PbY 154	80 03 17.05	−0.00288	0.00374	−0.0353	0.41	V		14.7 (W)	V_0 Hyperion = 14.2
SATU 29.11	80 02 29.18	0.00826	−0.000799	−0.0219	1.58	NF	15.3 (E)	15.5 (E) 15.7 (W)	B_0 Mimas = 13.5
SATU 18.01 B	80 03 18.04	−0.00351	0.00400	−0.0460	0.51	NF	14.7 (E)	15.2 (E) 14.8 (W)	B_0 Mimas = 13.5
SATU 18.07 A	80 03 18.10	−0.00358	0.00401	−0.0472	0.52	B	14.8 (E) 14.8 (W)	15.3 (E) 15.0 (W)	B_0 Tethys = 11.0

$\mu = \sin B$ and $\mu_0 = \sin B'$ where B and B' are the elevation of the Earth and the Sun respectively above the ring plane. α is the phase angle. NF, no filter; we used B mag to derive a value of the thickness. (E) and (W) refer to eastern and western ansae of ring. The saturnicentric distances of Sun and Earth were respectively 9.44 UA and 8.45 UA over this period.

and thus the limits of $m_v(\infty)$ and $m_B(\infty)$ when τ is infinite are easily obtained from these data. We found $m_v(\infty) = 6.32$ and $m_B(\infty) = 7.22$. These values have to be corrected from solar distance variation and from the so-called 'opposition effect', which occurs at small phase angle[14,15]. It is then possible to estimate the brightness of the illuminated side for an infinite optical thickness, with $B \simeq B' \simeq 26°$ at any phase angle $\alpha \leqslant 1°$. We assumed this brightness to be equal to b_E. In the simplest model (isotropic multiple scattering, no reflected light from Saturn's disk, see below), ϕ_S is approximatively a linear function of the product $B\times B'$, so that the extrapolation of $\phi = \phi_E + KBB'$ to $BB' = 0$ provides $z = \phi_E/b_E$ (Fig. 2).

Figure 2 shows the results of measurements at 20 arc s and 17 arc s from the centre of the planet (2.05 and 1.74 equatorial radii respectively). The ring seems to be systematically brighter in the outer part than in the inner part. We observe in particular a condensation of light around 21.8 arc s (ref. 10), which may be due to 'gaps' in A ring. On the other hand, the inner part is more opaque, and thus darker when transmitted light is observed. We used only the data at 17 arc s (Fig. 2*b*) for the determination of physical thickness z because the residual contribution of dark side at 20 arc s seems to remain significant, even near Earth's transit, on 12 March. The high value of z at $BB' = -0.0219$ (deg)^2 in Fig. 2*b* corresponds to an observation of 29 February (SATU 29.11 on Table 1) when the elevation B of the Earth was relatively large, and was probably due to the contribution of indirect illumination by Saturn's globe. This contribution is not significant at 20 arc s. These data give evidence of the reflected light from Saturn's globe, and cannot be used for a direct determination of the thickness. The vertical error bars on Fig. 2*a*, *b* are due to uncertainties in the microdensitometric measurements only. The horizontal bars are due to uncertainties in the values of B and B', due to an inaccuracy on Saturn's pole position. There are also uncertainties in satellite magnitudes (±0.15 mag), so that we derived from Fig. 2*b*:

$$z = 1.4 \pm 0.3 \text{ km}$$

Note that this is an estimate of the thickness of the model which may not be realistic. Actually, the notion of thickness itself is not evident as the rings may be warped, or surrounded by an 'atmosphere' of small particles, or by a belt of little satellites (two have already been discovered at 2.5 planetary radii whereas ring A terminates at 2.3 radii). Any of these phenomena would significantly contribute to the flux ϕ. These effects might explain the condensations observed in the rings (we notice as well an increase in brightness from the inner to the outer part of the rings on some plates taken on 12, 15 and 16 March)[10]. The cross-section of the edge is probably not rectangular as in Fig. 1, but rather curved or tapered. An upper limit of 1.7 km for the 'photometric apparent thickness' can be confidently claimed. We do not yet know precisely what the quantitative consequence of the edge effect is on the brightness of the edge that is the result of the escape of light through the sides of the ring (dashed rays on Fig. 1*b*). At any rate, the brightness due to single-scattering only is not affected by this edge-effect. The larger the asymmetry parameter g of the particle phase function, the weaker the spreading of light through both sides. The edge effect would be more important in blue light ($g = -0.328$) than in visible light ($g = -0.386$) (refs 14, 15). Preliminary estimates[9] seem to suggest that this effect could lead to an overestimate of b_E, and hence an underestimate of z, whose value could be 40% greater at most.

From the photographic plates of 1966, Focas and Dollfus[16] derived $z = 2.5 \pm 1.4$ km, whereas Lumme and Irvine[17,35] gave $0 < z < 2.8$ km. In 1980, better photometric accuracy of electronographic receptors, sophisticated computer image analysis and larger instruments have produced a more reliable photometry. Furthermore, the conditions of observation in 1980 were better than in 1966 as the crossing of rings plane in 1980 occurred closer to opposition when the elevation B' of the Sun was smaller (0.15° and 2.75° in 1980 and 1966 respectively) so that the side S facing the Earth was darker. There was no significant observation corresponding to BB' smaller than 0.014 in 1966 and thus the extrapolation was less accurate. In particular, we can claim that the 'photometric apparent thickness' is >1 km. The structure, the density and the particle albedo of the E ring are totally unknown. There are probably complicated structures and condensations in the equatorial plane as well as the perpendicular direction. Burns *et al.*[18] have shown that a warping of the Laplace plane can account for an apparent total 'thickness' as much as a few hundred metres. The observed brightness includes non-negligible contributions of the E ring, of large chunks and condensations and of the warping of the disk. In a disk made of a continuous distribution of particles of different sizes, as is the case for Saturn's rings, the contribution of the large particles to the 'apparent photometric thickness' predominates. Thus, observations at ring-plane passage do not necessarily suggest a 'thick' ring as it is often claimed[19] and are not necessarily in contradiction with a ring maintaining a thickness of only a few particles radii as suggested by the study of systems of inelastically colliding particles[20,21]. On the contrary, the 'theoretical thickness' can be larger than a few particles radii.

Thickness of a three-dimensional system of inelastically colliding particles

The controversy over the vertical structure of the rings (see for examples refs 22, 23), whether it is a single layer in thickness, a few layers or many layers thick, is not merely an academic exercise; it has implications for the dynamics and evolution of the rings as well as for the mean sizes and surface topographies of the ring particles[12,20,21].

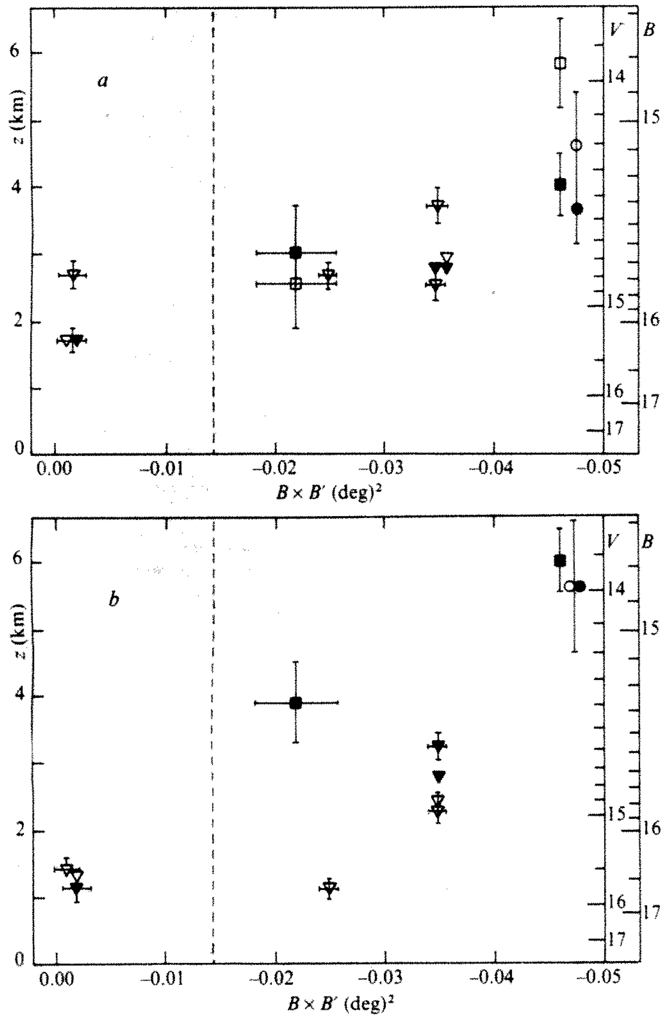


Fig. 2 Thickness ϕ_e/b_e (km) as a function of the date of observation (or of the product of the angle BB' (deg^2)). The quantity ϕ_e/b_e expressed in the text in arc s, has been converted into kilometres. On each plate two scans were made, when possible, at 20 arc s from the centre of the planet (East and West sides), and two at 17 arc s. We have so far reduced plates in blue and visible lights; the scales on the right provide the mag/arc s of the rings in these two lights. The broken line indicates the lowest value of BB' for the observations in 1966. Without data represented by points on the left, the extrapolation to obtain the apparent thickness is much more difficult.

Side / Colour	East	West
Visible	▼	▽
Blue	●	○
Total	■	□

a, Data corresponding to a scan at 20″. *b*, Data corresponding to a scan at 17″.

It has been shown[20,21,24] that after a very fast flattening of the order of 20 collisions per particle, a three-dimensional system of inelastically colliding particles reaches a quasi-equilibrium state in which the thickness of the newly formed disk is finite and in which collisions still occur. The combined effect of differential rotation and inelastic collisions, makes the disk spread very slowly, particles moving both inwards and outwards carrying out some angular momentum. For Saturn's rings, the time scale of flattening is of the order of a few weeks, the time scale of the quasi-equilibrium state is of the order of 10^{14} yr and the system reaches a third hypothetical collisionless state in a time scale of the order of 10^{21} yr (refs 20, 21). Although this result disagrees strongly with the time scale obtained by Jeffreys[25], his paper contains an erroneous calculation (see ref. 24, and M. Hénon, personal communication).

To reconcile the 'many-particle thick' conditions apparently required by recent models of ring photometric properties, which are derived from a variety of optical (opposition effect) and radio observations[23,26,27], Burns, Cuzzi *et al.*[18,28,29] have considered several mechanisms as possible sources of energy for continually thickening the rings such as: gravitational perturbations by Saturnian satellites, the Sun, and the oblateness of Saturn; meteoroid bombardments; radiation pressure, and mutual gravitational scattering of particles.

We have recently studied simple models of three-dimensional systems of inelastically colliding particles in which the attraction between particles has been neglected (and so particle orbits are Keplerian around a central mass). In these models particles have the same masses and radii and the loss of energy during a collision does not depend on the relative velocity of the two colliding particles. When a large part of the orbital energy is lost after a collision, an instability phenomenon appears. For very inelastic collisions and contrary to the behaviour of the first models[21] (Fig. 3, zone B), the disk is completely flattened and the system becomes a two-dimensional configuration in which collisions continue to occur (Fig. 3, zone A). A sharp transition L_1 separates zones A and B. A linear stability analysis of the observed behaviour has been made: differential rotation constantly feeds energy into horizontal motions and thus maintains a finite dispersion of horizontal velocities, whilst vertical motions are only a by-product of the horizontal activity and do not necessarily arise[30]. A linear stability analysis of a flat disk of colliding particles gives a condition for stability:

$$\gamma(k-k')^2 < 4(1-k'^{-2}) \qquad (1)$$

where $\gamma = \langle V_r^2/\omega^2 r^2\rangle$ v, is the relative velocity of two particles, r the radius of one particle, ω the angular velocity, $1+k$ and $1-k'$

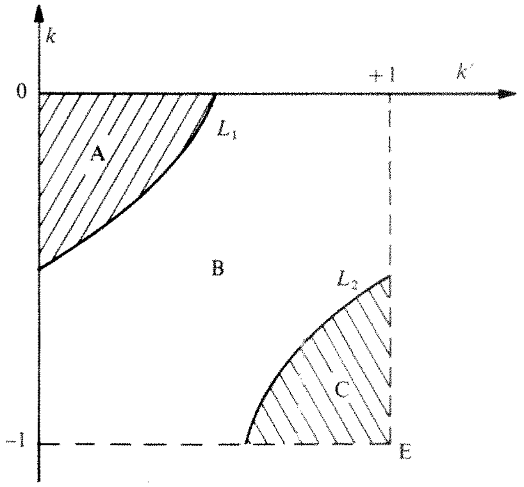


Fig. 3 The flattening of the system is a function of the rebound coefficients k and k'. In zone A, the system is completely flattened. In zone B the thickness of the disk decreases towards a value of the order of a few times the size of a particle. In zone C, the thickness of the disk increases. There exist two sharp limits L_1 and L_2 between the three zones, A, B, and C.

the relative loss of velocity after one collision in the normal and the grazing direction respectively. $(k, k') = (0, 0)$ corresponds to a complete inelastic collision and $(k, k') = (-1+1)$ to an elastic collision. The evolution of models with velocity dependance of k, k' is essentially the same[31]. There is a second sharp transition L_2 between zone B and zone C (Fig. 3). In zone B, the thickness of the disk decreases and reaches a value of the order of r/R, where r is the radius of a particle and R some characteristic dimension of the system. In zone C, the thickness of the disk increases and the system becomes a true three-dimensional system. On the one hand, keplerian motion tends to establish an anisotropic distribution of velocities with a velocity twice as large in the radial direction than in the transverse direction. On the other hand, collisions tend to establish an isotropic distribution of velocities. If the degree of inelasticity is larger (Fig. 3, zones A and B), a smaller quantity of energy is extracted from the ordered motion and the orbits become more and more circular and co-planar. A crude calculation of the equilibrium between these two phenomena gives the equation:

$$L_2\colon k^2 + k'^2 = 1.28 \qquad (2)$$

(For L_2 see ref. 31.) Zone C probably does not correspond to the Saturn case, unless nonlinear effects damp the vertical motions when the system becomes very thick.

In a real system, most collisions would be between particles of very unequal size[32]. We shall give here only those results concerning the mutual evolution of two different sized particles. Generally, the two groups of particles evolve in essentially the same way. There is no real equipartition of energy, but rather a kind of separation. The mean inclination and the mean eccentricity of the large particles is smaller than the mean inclination and the mean eccentricity of the small particles (Fig. 4). The evolution is essentially similar to that of the first models. The more elastic the collisions, the bigger the separation; the larger the mass ratio, the bigger the separation; there would not seem to be any separation in the plane of the disk. One can see numerically and by a crude calculation[33] that the mean velocity $\langle v_l \rangle$ of the large particles is proportional to the mean velocity $\langle v_s \rangle$ of the small particles:

$$\langle v_s \rangle \propto \frac{1}{(1-k^2)^{1/2}} \langle v_l \rangle \qquad (3)$$

Generally the massive particles evolve faster than the lighter ones. The relation ($ar^2 < 3 \times 10^{10}$ yr m^2) between the size r of the particles and the age a of the system, shown—for systems of identical particles[21], remains valid for each component individually. To the extent that Saturn's rings may be modelled in this way, then if the rings are as old as the Solar System, the maximum size of the particles is 2.5 m. A system made essentially of large particles, would presumably have been created much later. Cuzzi *et al.*[29] have shown that the ring systems attains a steady state in which gravitational scattering (not considered here) by the large particles impacts random velocity components to the small particles while inelastic collisions (not quantitatively treated in the Cuzzi's model) descrease these random velocities.

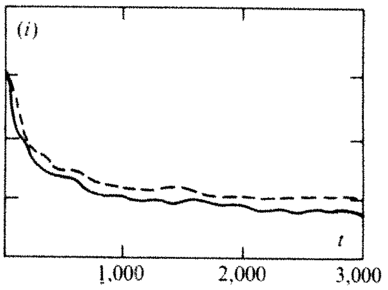


Fig. 4 Variations as a function of time of the mean inclination of the particles (which is a good measure of the flattening of the system). The results for large particles are shown by a solid line and those for small particles are given by the dashed line. The particles have the same density.

We have studied systems in which a satellite orbits around the primary on an orbit inclined with respect to a disk of colliding particles. The influence of the satellite on the thickness is negligible. Detailed results on small changes in the mean plane's location due to the action of the satellite are given elsewhere[34].

Conclusion

The main results of our numerical simulation are:

(1) In the case of almost elastic collisions, a disk of colliding particles can expand vertically to become a three-dimensional system.

(2) When particles of different masses and sizes collide there is no real equipartition of energy, but rather a kind of separation between the different populations.

(3) Satellite perturbations are not very efficient in thickening a disk of colliding particles.

The 'local' thickness of the rings is probably of the order of a few particles radii or some tens of metres[21]. Unless some as yet not understood mechanism reduces the vertical motions, the phenomenon observed in zone C of Fig. 3 is apparently not present in Saturn's rings, as it would lead to a much larger thickness. Contrary to what one could deduce from a relative velocity of the order of only 10^{-2} cm s^{-1} for 2-m sized colliding particles in Saturn's rings, there is enough energy lost during the collisions. If not the rings would be much thicker. On a large scale, the rings are probably slightly warped[18] and are possibly surrounded by an atmosphere[19] and by small satellites or large chunks of matter just outside (as discovered in March 1980)[4,5] or even inside the rings. This would produce on average an apparent edge-on thickness of several hundred metres and would make the true physical thickness (a local vertical scale height of the rings) impossible to determine from ring-passage observations from the Earth or from fly-bys.

We thank M. Hénon for helpful advice on the theoretical studies; J. Lecacheux, P. Laques, R. Despiau, G. Wlérick, B. Servan, D. Michet, L. Renard, and A. Bijaoui for observations and assistance, and A. Cazenave, J. Burns and R. Greenberg for stimulating discussions. Much of this work was supported by the RCP 544 'Anneaux' and the ATP 'Planétologie' of the CNRS.

Received 1 September; accepted 18 November 1980.

1. Elliott, J. L., Dunham, E., Wasserman, L. H., Millis, R. L. & Churms, J. *Astr. J.* **83**, 980–992 (1978).
2. Owen, T. *et al. Nature* **281**, 442–446 (1979).
3. *Science* **207**, 400–453 (1980).
4. Lecacheux, J., Laques, P., Vapillon, L., Auge, A. & Despiau, R. *Icarus* **43**, 112–115 (1980).
5. Reitsema, H. J., Smith, B. A. & Larson, S. M. *Icarus* **43**, 116–119 (1980).
6. Dollfus, A. & Brunier, S. *C.r. hebd. Séanc. Sci., Paris* **290**, 261–263 (1980).
7. *IAU Circ.* Nos 3426, 3430, 3454, 3456, 3457, 3463, 3466, 3483, 3484, 3496 (1980).
8. Fort, B. *et al. Icarus* (submitted).
9. Sicardy, B., Lecacheux, J., Laques, P., Despiau, R. & Augé, A., *Astr. Astrophys.* (submitted).
10. Sicardy, B. *et al. C.r. hebd. Séanc. Acad. Sci. Paris* **291**, 1–6 (1980).
11. Veverka, J., in *Planetary Satellites* (ed. Burns, J.) 171 (University of Arizona Press, 1977).
12. Bobrov, M. S. *Soviet Astr., A.J.* **16**, 348–354 (1972).
13. Cook, A. F., Franklin, F. A. & Palluconi, F. D. *Icarus* **18**, 317–337 (1973).
14. Lumme, K. & Irvine, W. M. *Astr. J.* **81**, 865–893 (1976).
15. Lumme, K. & Reitsema, H. J. *Icarus* **33**, 288–300 (1978).
16. Focas, J. H. & Dollfus, A. *Astr. Astrophys.* **2**, 251–265 (1969); *C.r. hebd. Séanc. Acad. Sci., Paris* **268**, 100 (1969).
17. Lumme, K. & Irvine, W. H. *Astr. Astrophys.* **71**, 123–130 (1979).
18. Burns, J. A., Hamill, P., Cuzzi, J. N. & Durisen, R. H. *Astr. J.* **84**, 1783–1801 (1979).
19. Bobrov, M. S. in *Surfaces and Interiors of Planets and Satellites* (ed. Dollfus, A.) 438 (Academic, New York, 1970).
20. Brahic, A. *Icarus* **25**, 452–458 (1975).
21. Brahic, A. *Astr. Astrophys.* **54**, 895–907 (1977).
22. Cook, A. F. & Franklin, F. A. in *Planetary Satellites* (ed. Burns, J. A.) 412 (University of Arizona Press, 1977).
23. Pollack, J. B. *Space Sci. Rev.* **18**, 3–93 (1975).
24. Goldreich, P. & Tremaine, S. *Icarus* **34**, 227–239 (1978).
25. Jeffreys, H. *Mon. Not. R. astr. Soc.* **107**, 263–267 (1947).
26. Cuzzi, J. N. & Pollack, J. B. *Icarus* **33**, 233–262 (1978).
27. Kawata, Y. & Irvine, W. H. in *Exploration of the Planetary System* (eds Woszczyk, A. & Iwaniszewska, C.) 441 (Reidel, Dordrecht, 1974).
28. Cuzzi, J. N., Burns, J. A., Durisen, R. H. & Hamill, P. *Nature* **281**, 202–204 (1979).
29. Cuzzi, J. N., Durisen, R. H., Burns, J. A. & Hamill, P. *Icarus* **38**, 54–68 (1979).
30. Brahic, A. & Hénon, M. *Astr. Astrophys.* **59**, 1–7 (1977).
31. Brahic, A. *Astr. Astrophys.* (submitted).
32. Greenberg, R., Walker, J. F., Hartmann, W. K. & Chapman, C. R. *Icarus* **35**, 1–26 (1978).
33. Brahic, A. *Astr. Astrophys.* (submitted).
34. Brahic, A. *Astr. Astrophys.* (submitted).
35. Dollfus, A. *Astr. Astrophys.* **75**, 204–206 (1979).

New data base for climate studies

Michael Matson & Donald R. Wiesnet

NOAA/National Earth Satellite Service, Washington DC 20233

Until the advent of satellite data it had been impractical to measure continental snow cover. Beginning with the ESSA 3 satellite in late 1966, and continuing to the present NOAA and GOES series satellites, Northern Hemisphere continental snow cover has been routinely mapped by the National Oceanic and Atmospheric Administration (NOAA). This mapping has produced a reliable, timely and relatively unbiased continental snow-cover data base for use in climate studies.

IN recent years the impact of global snow and ice cover on climatic processes has received increased attention[1-5]. Kukla[6] stated that "to understand properly the mechanisms of weather and climate variability we need to understand the variability of snow and pack-ice fields". The Global Atmospheric Research Program (GARP) has recommended that sensitivity studies of the climatic effects of global and regional anomalies of snow and ice cover be performed with various types of general circulation models[7]. Specifically, continental snow cover has important climatic impacts on the radiation balance, surface and air temperature, cloudiness, soil moisture, water storage and precipitation. Continental snow cover can also have an impact on energy requirements, transportation systems and food supplies.

Because of the climatic–economic importance of continental snow cover, GARP[7], the Committee on Space Research (COSPAR)[8], and the United States Climate Program Plan[9], have recommended that global monitoring of the extent of snow cover be included in both short- and long-term observational programmes. Until 1966, however, the monitoring of the areal extent of continental snow cover was limited to point measurements and extrapolation between these often widely separated points. Mountainous and sparsely inhabited areas were poorly represented in this type of monitoring. In November 1966 the first satellite-derived Northern Hemisphere Weekly Snow and Ice Cover Chart was produced by the National Environmental Satellite Service (NESS) of NOAA. Since then, continental snow-cover extent has been monitored on a weekly basis by satellites. In this article we demonstrate the ability of satellites to monitor an important climatic parameter and discuss the 14-yr satellite record of continental snow cover.

The NOAA/NESS satellite capability

The NOAA/NESS satellites and sensors used in mapping Northern Hemisphere continental snow cover are listed in Table 1. Note that the high-resolution VHRR, AVHRR and VISSR data are primarily available only for North America. Table 1 shows that satellite data were available at all times from late 1966 to the present for monitoring continental snow cover. Of course, the snow-cover data have come from a variety of satellites and sensors, all subject to instrumental variation, degradation and drift. The impact of these factors on mapping continental snow cover, however, has been negligible.

Table 1 NOAA/NESS satellites and sensors used in mapping Northern Hemisphere snow and ice cover

Satellite	Sensor*	Spectral band (μm)	Subpoint resolution (km)	Period of operations
ESSA 3	AVCS	0.5–0.75	3.7	2 October 1966–9 October 1968
ESSA 4	APT	0.5–0.75	3.7	26 January 1967–6 December 1967
ESSA 7	AVCS	0.5–0.75	3.7	15 August 1968–19 July 1969
ESSA 8	APT	0.5–0.75	3.7	15 December 1968–12 March 1976
ESSA 9	AVCS	0.5–0.75	3.7	26 February 1969–15 December 1973
ITOS 1	AVCS	0.5–0.75	3.7	23 January 1970–17 June 1971
	APT	0.5–0.75	3.7	
	SR	0.52–0.73	3.7	
		10.5–12.5	7.4	
NOAA 2	VHRR	0.6–0.75	1–1.9	15 October 1972–30 January 1975
		10.5–12.5	1–1.9	
	SR	0.52–0.73	3.7	
		0.50–0.94	3.7	
		10.5–12.5	7.4	
NOAA 3	VHRR	0.6–0.75	1–1.9	6 October 1973–17 December 1974
	SR	0.52–0.73	3.7	
		10.5–12.5	7.4	
NOAA 4	VHRR	Same as NOAA-3		15 November 1974–15 September 1976
	SR	Same as NOAA-3		
NOAA 5	VHRR	0.6–0.75	1–1.9	29 July 1976–1 March 1979
		10.5–12.5	1–1.9	
	SR	0.52–0.73	3.7	
		0.50–0.94	3.7	
		10.5–12.5	7.4	
TIROS N	AVHRR	0.55–0.68	1	6 November 1978–20 January 1980
		0.725–1.10	1	
		3.55–3.93	1	
		10.5–11.5	1	
NOAA 6	AVHRR	Same as TIROS-N		16 July 1979–Present
GOES series	VISSR	0.55–0.70	1–7.4	17 May 1975–Present
		10.5–12.5	7.4–14.8	

* Cameras and sensors: AVCS, advanced vidicon camera system; APT, automatic picture transmission; SR, scanning radiometer; VHRR, very high resolution radiometer; AVHRR, advanced very high resolution radiometer; VISSR, visible and IR spin scan radiometer.

0028–0836/81/060451—06$01.00

The NOAA/NESS Northern Hemisphere snow-cover data base

Since November 1966, using the visual band sensors listed in Table 1, the Synoptic Analysis Branch (SAB) of NESS has prepared the Northern Hemisphere Weekly Snow and Ice Cover Chart (Fig. 1). These charts, at a scale of 1:62,500,000, show the areal extent and brightness of continental snow cover but do not indicate snow depth. The preparation of the charts and the definition of terms used in the analysis have been described in detail in an unpublished report by SAB; we give here only brief details.

Early in the week the analyst, a satellite meteorologist, makes a pencil trace of the previous week's chart. Daily he collects all of the visual band data received during the previous day, compares snow and ice cover interpreted from the imagery with snow cover on the previous week's chart, and makes changes. Each area of the new chart is therefore the latest cloud-free snow observation of that particular area of the world. If an area is cloud covered for several days, the analysis of snow cover for

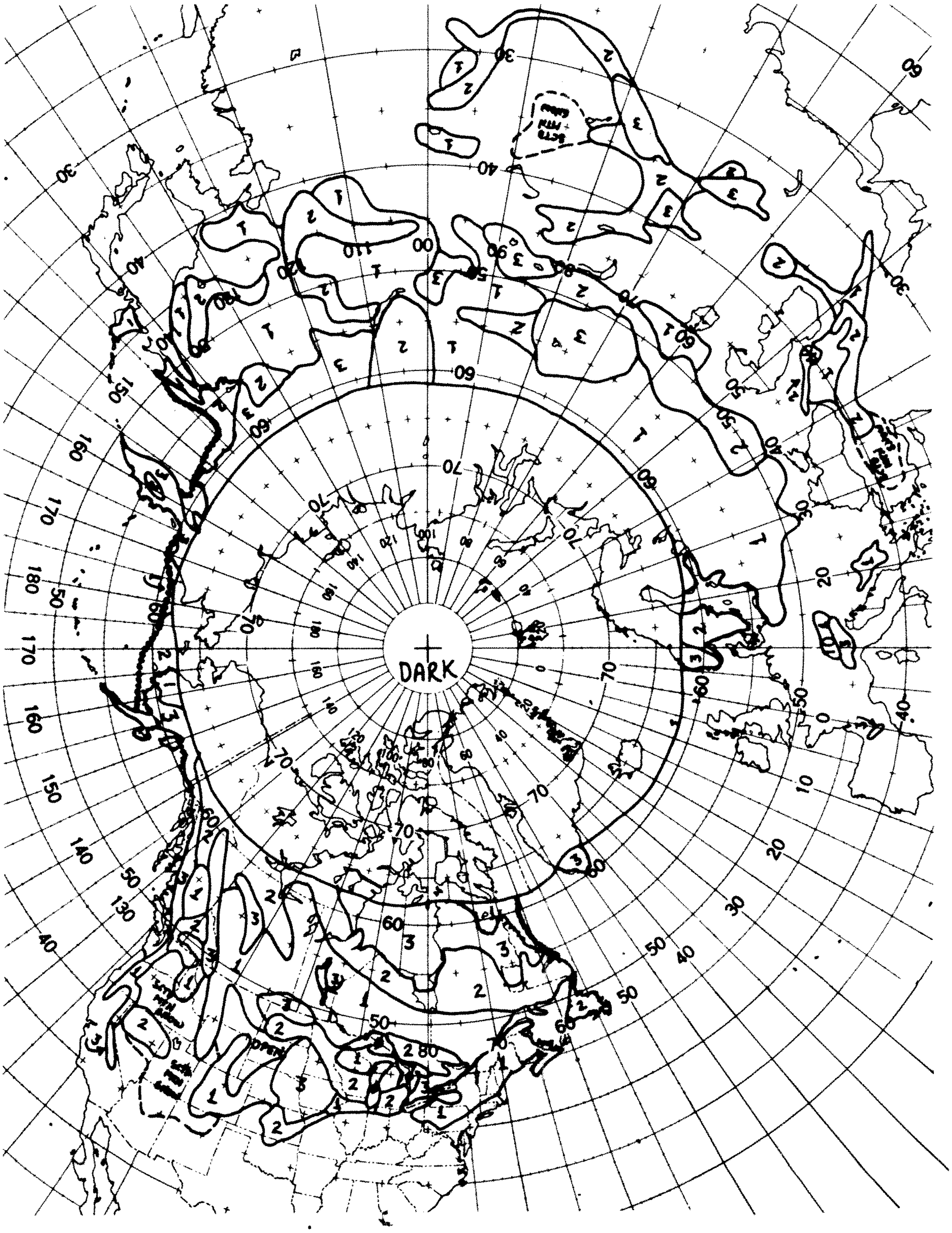


Fig. 1 Typical snow and ice chart of the Northern Hemisphere for the period 29 December 1974 to 4 January 1975 (scale 1:62,500,000), prepared by the Synoptic Analysis Branch, NOAA/NESS. Reflectivities: 1, lowest; 2, moderate; 3, highest. —, snow; ○, ice. Note the 'dark' area where visible data cannot be collected during the polar winter.

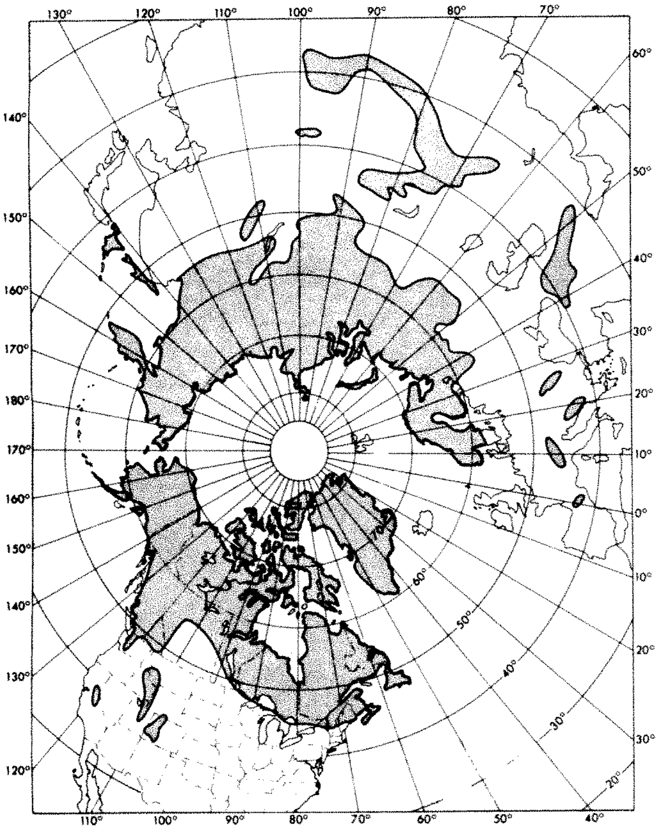
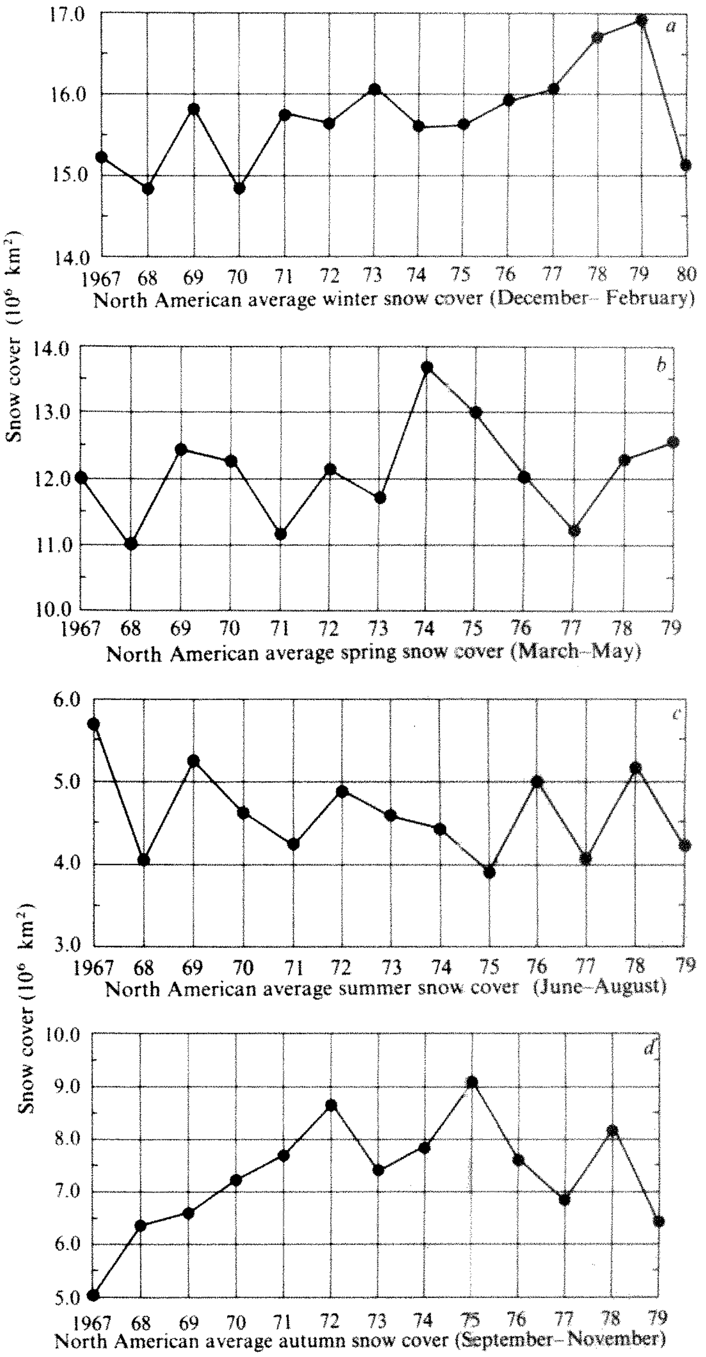


Fig. 2 Monthly mean snow cover chart of the Northern Hemisphere for April 1978 on a stereopolar projection at a ~1:100,000,000 scale. This chart was prepared from the weekly charts.

that area will necessarily be several days old. If the area remains cloud covered for the entire week, the previous week's analysis is carried over to the new chart. Finally, the analyst compares his chart with the surface synoptic reports to check for errors or to confirm drastic changes in snow cover, as can result from a new snowfall or rapid snow melt. At the end of the week the new chart is rendered in final form and transmitted to the user community. Solid lines on the chart indicate boundaries between areas of substantially differing brightness (reflectance) of the snow field. The chart is divided into three categories of brightness: (1) lowest reflectivity; (2) moderate reflectivity; and (3) highest reflectivity. These categories are subjective and are chosen by the analyst. Guidelines aiding the analyst in assigning these reflectivity classes include the following. First, the designated reflectivity of a given snow area is the average reflectivity within that boundary, that is, a category (2) could be the average of a small area of category (3) and a small area of category (1), or the entire area could be a category (2) reflectivity. Second, the reflectivity of categories is relative to their surroundings at a particular time of year. For example, although it is much less bright in the winter due to reduced solar illumination, the Greenland icecap is labelled category (3) in both summer and winter. It is, however, much brighter than its surroundings even at this time. Third, when snow-covered mountain tops are seen in large open areas that average less than category (1), the area is labelled as scattered mountain snow. Fourth, when the distinction in reflectivity between small areas of land and water is great, such as over Arctic Canada in early summer, the entire area is labelled 3 over water and 1 over land.

In winter much of the far northern latitudes lies in total darkness or near darkness. Consequently, the visual-band sensors on the satellite are inoperative, and areas around the pole are shown as 'dark'. At the winter solstice the snow and ice chart is labelled 'dark' north of about 55 °N.

The quality of the weekly charts is affected by several factors: The charts are based on interpretations by a variety of observers,



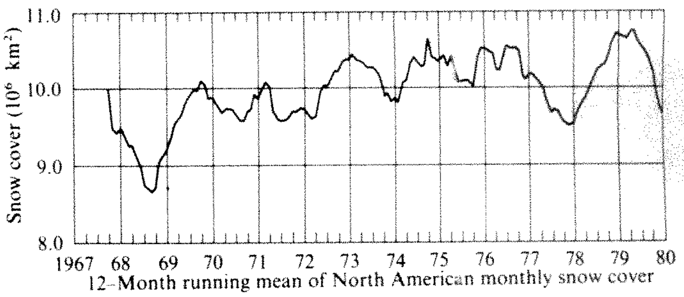
North American average winter snow cover (December–February)

North American average spring snow cover (March–May)

North American average summer snow cover (June–August)

North American average autumn snow cover (September–November)

Fig. 3 *a*, Average North American winter snow cover for the period 1967–1980. Calendar year notation is such that 79 represents the winter of 1978–79. *b*, Average North American spring snow cover for the period 1967–79. *c*, Average North American summer snow cover for the period 1967–79. *d*, Average North American autumn snow cover for the period 1967–79.



12-Month running mean of North American monthly snow cover

Fig. 4 Twelve-month running mean of North American monthly mean snow cover. The data are plotted on the last month of the 12-month period.

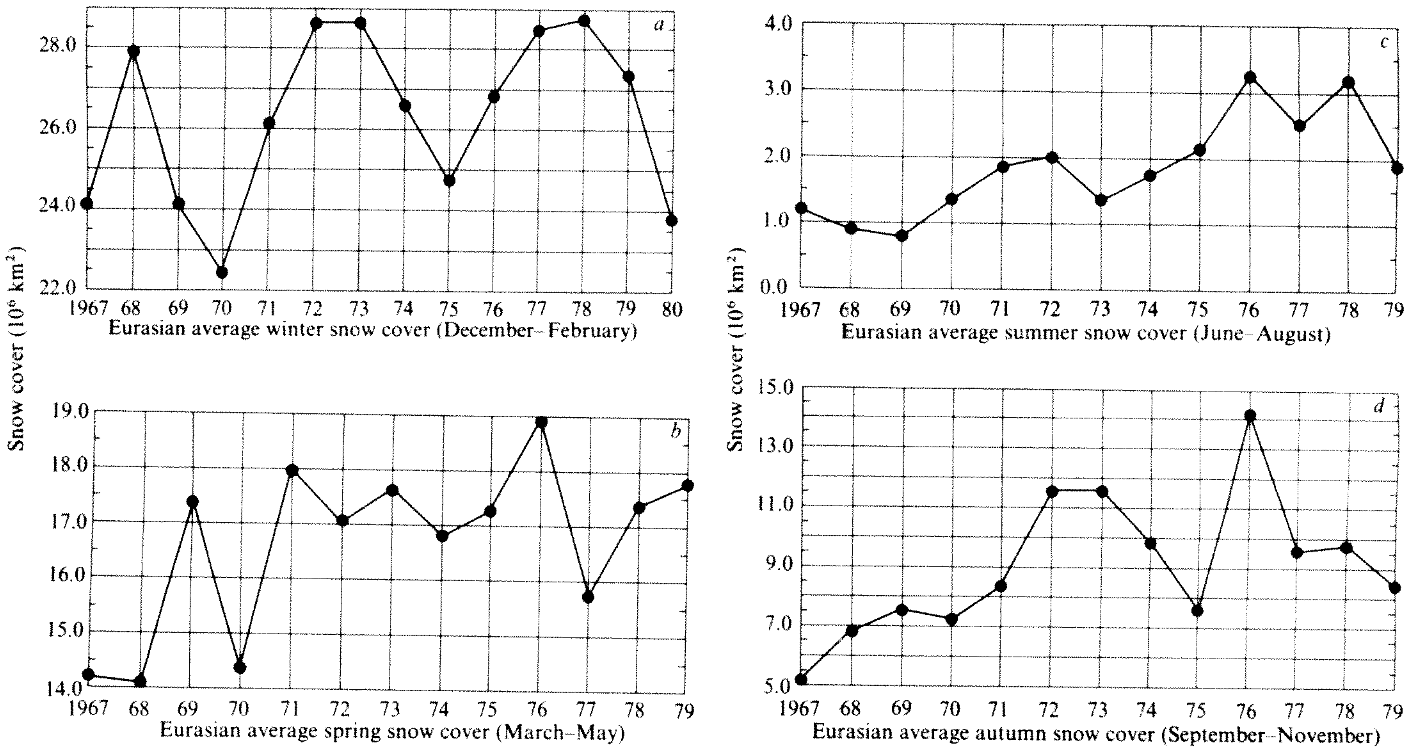


Fig. 5 *a*, Average Eurasian winter snow cover for the period 1967–80. Calendar year notation is such that 79 represents the winter of 1978–79. b, Average Eurasian spring snow cover for the period 1967–79. *c*, Average Eurasian summer snow cover for the period 1967–79. *d*, Average Eurasian autumn snow cover for the period 1967–79.

thus operator bias is undoubtedly present; the scale of the charts limits the precision of snow line positioning to about 20–50 km (ref. 10); Himalayan snow cover was not consistently mapped during the 1966–74 period; the charts show three levels of increasing improvement in consistency and detail—1966–70, 1971–73, and 1974 to the present. Kukla and Robinson[11] analysed the accuracy of the NOAA/NESS weekly charts by comparing them with an independent daily analysis of weekly snow cover based on satellite data and ground station reports for the United States. The analysis covered periods for November 1975, February 1977 and February 1978. After measuring the areal extent of snow cover derived from both the NOAA/NESS and independent analysis, they concluded that: the areal differences tend to be smaller in the more recent charts; the average areal differences are usually less than 10%; the accuracy of the NOAA/NESS weekly charts is adequate for climate-related studies on a continental or hemispheric scale.

Monthly mean snow-cover charts (Fig. 2) have been constructed from the weekly charts for all months from November 1966. The winter (December to March) monthly snow-cover maps from 1966 to 1976 have been published[12] and a digitized atlas of all of the monthly maps is being prepared. The monthly charts are constructed by deriving a subjective average of the NOAA/NESS weekly chart boundaries of each month. The areal extent of continental snow cover within this average monthly snow-cover boundary is then measured and recorded. All continental areas within the 'dark' zone are assumed to be completely snow-covered and are measured as such. The snow-cover values are corrected to represent true surface area coverage. The error in the value of areal snow cover owing to the analogue averaging of weekly data to derive the monthly snow line is approximately 2–5% and randomly distributed[2].

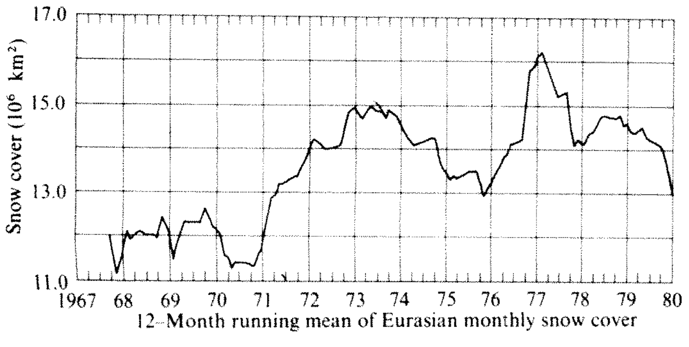


Fig. 6 Twelve-month running mean of Eurasian monthly mean snow cover. The data are plotted on the last month of the 12-month period.

A look at the record

Figure 3 shows a graphical record of North American seasonal snow cover from 1967 to 1980. The average winter snow cover (Fig. 3*a*) rose steadily from 1975 to 1979—an 8% increase when 1979 is compared with 1975. The winters of 1977, 1978 and 1979 are discussed in more detail elsewhere[13–15]. During the severe North American winter of 1978 nearly 70% of the continent was snow-covered. Although characterized by marked fluctuations, spring (Fig. 3*b*) and summer (Fig. 3*c*) average snow cover has been around 12.1×10^6 km^2 and 4.6×10^6 km^2, respectively. Autumn average snow cover (Fig. 3*d*) shows two distinct periods: a period of increase from 1967 to 1972 and a period of levelling off with sharp annual fluctuations from 1972 to 1978. Figure 4 is a 12-month running mean (to remove the annual cycle) of North American monthly snow cover and it shows a slight tend towards increasing snow cover for the continent.

As Eurasia is a larger continent than North America, Eurasian seasonal snow-cover fluctuations exhibit a wider range of values, as revealed by Fig. 5. Winter snow cover (Fig. 5*a*) has an interesting cyclic pattern but the data set is too short to confirm the apparent repeatability. Very large areas of the continent were covered by snow during the winters of 1972, 1973, 1977 and 1978. At these times over 50% of the Eurasian land mass was snow-covered. Spring snow cover (Fig. 5*b*) varies markedly, showing especially large annual changes from 1968 to

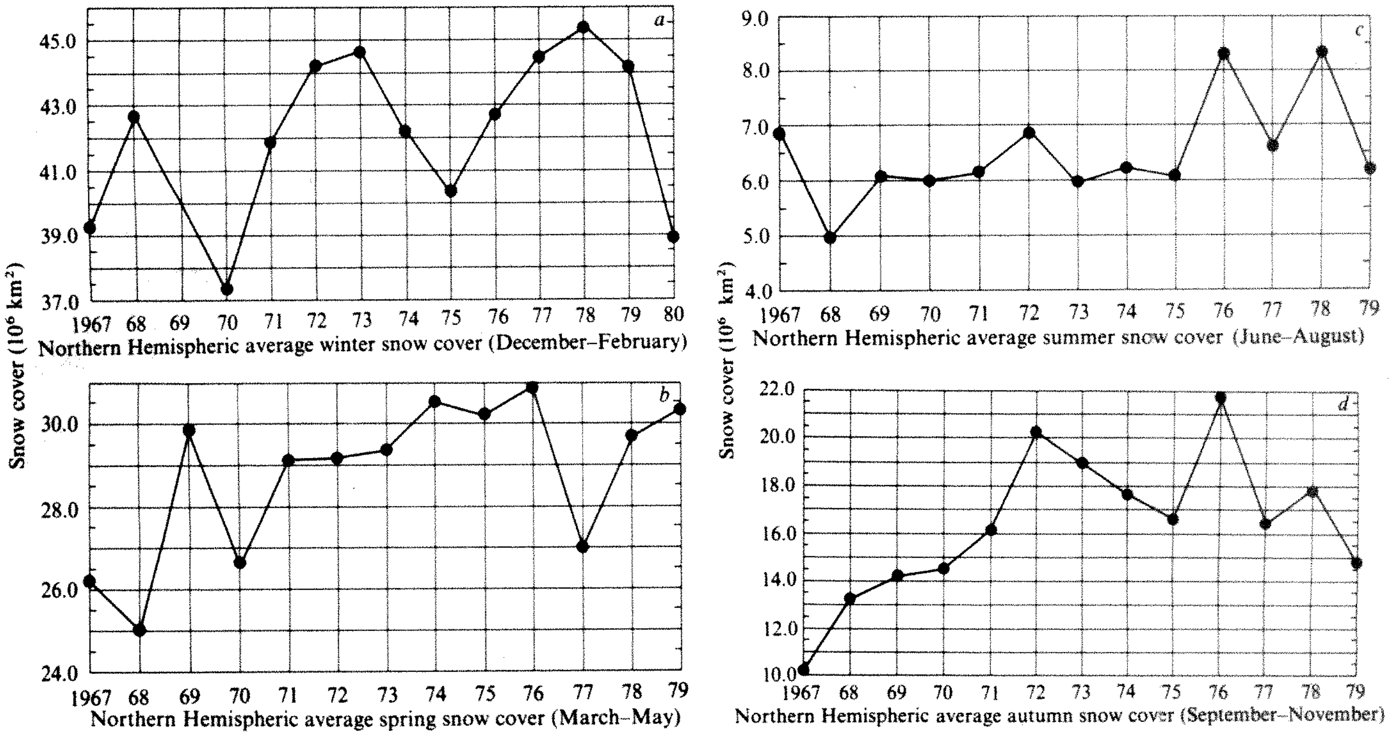


Fig. 7 *a*, Average Northern Hemisphere winter snow cover for the period 1967–80. Calendar year notation is such that 79 represents the winter of 1978–79. *b*, Average Northern Hemisphere spring snow cover for the period 1967–79. *c*, Average Northern Hemisphere summer snow cover for the period 1967–79. *d*, Average Northern Hemisphere autumn snow cover for the period 1967–79.

1971 and 1975 to 1978. Although Eurasian summer snow cover (Fig. 5*c*) shows an increasing trend, the reality of that trend is dubious because of the previously mentioned inconsistency in mapping Himalayan snow cover during the 1966–74 period. The average of the summer snow-cover values from 1975 to 1979, 2.6×10^6 km^2, is probably a more realistic value than the earlier summer values. Autumn snow cover (Fig. 5*d*) also shows an increasing trend over the whole satellite record with a particularly sharp increase in 1976. This increase was in a large part due to a satellite-based record high October snow cover of 19.7×10^6 km^2 compared with a 10-yr (1967–76) average of 9.6×10^6 km^2. Most of this record snow cover occurred from 11 to 17 October when the snow line advanced from 60 to 50 °N in an area longitudinally bordered by 30 °E and 80 °E. This area represents more than 3×10^6 km^2 of snow cover, all in a 1-week period. The snow cover certainly contributed to the high (45%) satellite-derived October zonal albedo value for the area 0–117.5 °E and 60–50 °N (ref. 16). Once established the snow cover persisted throughout the autumn and into the winter. The effect of this anomalous and sudden snow cover on the radiation balance and atmospheric circulation would be interesting for researchers concerned with climate feedback mechanisms.

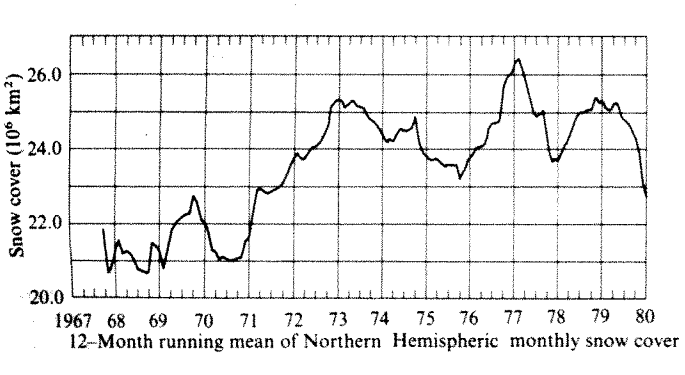


Fig. 8 Twelve-month running mean of Northern Hemisphere monthly mean snow cover. The data are plotted on the last month of the 12-month period.

The Eurasian 12-month running mean of monthly snow cover is shown in Fig. 6. During the satellite record Eurasia has experienced two periods of large increases in snow-cover extent, one from 1971 to 1973, the other from 1976 to 1977. The running mean for Eurasia also shows an overall upward trend.

Except for summer, the proportionally larger variations in Eurasian snow cover dominate the North American variations. The result is that trends in the Northern Hemisphere and Eurasian data for the winter, spring and autumn are similar (Fig. 7*a*, *b*, *d*). During the summer season most of the Northern Hemisphere snow cover is in Greenland and the Canadian archipelago, as shown by the similarities in the graphs of Fig. 3*c* and Fig. 7*c*. The 12-month running mean of Northern Hemisphere snow cover (Fig. 8) is also dominated by the Eurasian snow cover (Fig. 6), both showing similar trends and fluctuations. From these similarities one may conclude that in any study of hemispheric climatic feedback mechanisms involving snow cover, Eurasia has a greater feedback potential than North America.

Conclusions

The present satellite-derived data provide a rapid and timely synthesis of global snow-cover information to assess current conditions. This accumulation of a reliable and relatively unbiased continental snow-cover climatic data base has application in seasonal weather forecasting and in global and regional modelling. As the satellite snow-cover record is extended temporally, it should provide a new understanding of the dynamics of global and regional climate.

We thank Paul McClain for his review, Dennis Haller of SAB for the unpublished description of the construction of the Northern Hemisphere Weekly Snow and Ice Cover Charts, George Kukla for his continued support, Craig Berg for his analysis of the snow cover data during 1978–79, and Mrs Olivia Smith for assistance.

Received 14 July; accepted 13 November 1980.

1. Kukla, G. J. & Kukla, H. J. *Science* **183,** 709–714 (1974).
2. Wiesnet, D. R. & Matson, M. *Mon. Weath. Rev.* **104,** 828–835 (1976).
3. Hahn, D. G. & Shukla, J. *J. atmos. Sci.* **33,** 2461–2462 (1976).
4. Robock, A. *J. atmos. Sci.* **35,** 1111–1122 (1978).
5. Kukla, G. J. *et al. Nature* **270,** 573–580 (1977).
6. Kukla, G. J. in *Climatic Change* (ed. Gribbin, J.) 114–130 (Cambridge University Press, 1978).
7. GARP *The Physical Basis of Climate and Climate Modelling* 265 (GARP Publ. Ser. No. 16, 1975).
8. COSPAR *Status of Satellite Observing Possilbilities for Studies of Climate Physical Processes* 113 (NCAR, Boulder 1978).
9. *A United States Climate Program Plan,* 81 (NOAA, Washington DC, 1977).
10. Kukla, G. J. *Proc. Symp. on Meteorological Observations from Space: Their Contribution to the First GARP Global Experiment* 110–115 (NCAR, Boulder, 1976).
11. Kukla, G. J. & Robinson, D. *Glaciological Data Rep. GD-5,* 91–98 (NOAA, Washington DC, 1979).
12. Matson, M. *NOAA Tech. Mem. NESS* **84,** 24 (1977).
13. Matson, M. *Proc. NOAA Climate Diagnostics Workshop,* 1010–1016 (NOAA, Washington DC, 1977).
14. Wiesnet, D. R. & Matson, M. *Mon. Weath. Rev.* **107,** 928–933 (1979).
15. Wiesnet, D. R. & Berg, C. P. *Proc. 5th A. William T. Pecora int. Symp. on Remote Sensing* (in the press).
16. Winston, J. S. *et al. Earth–Atmospheric Radiation Budget Analyses Derived from NOAA Satellite Data June 1974–February 1978* Vol. 2 (NOAA, Washington DC, 1979).

Identification of a glycophorin-like molecule at the cell surface of rat thymocytes

William R. A. Brown, A. Neil Barclay, Christopher A. Sunderland* & Alan F. Williams

MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK

There are three predominant glycoproteins in the rat thymocyte plasma membrane. Two of these have carbohydrate compositions that are characteristic of structures N*-glycosidically linked to protein. The other glycoprotein is very different, having about 20* O*-glycosidically linked carbohydrate units per 100 amino acids. It has similarities to the major sialoglycoprotein of human erythrocytes.*

CELL-SURFACE glycoproteins with large amounts of carbohydrate are found on most eukaryotic cell types. Although little is known of the structure and functions of these molecules, it is widely thought that they may be involved in determining interactions between cells[1,2]. We have now completed the initial characterization of the predominant glycoproteins of rat lymphoid cells.

Much of the membrane carbohydrate on rodent lymphoid cells is found on a small number of cell-surface glycoproteins. These can be identified on intact cells by labelling sialic acid residues with ^{3}H-borohydride after mild periodate oxidation[3]. In the rat and mouse the glycoproteins so labelled are very similar in apparent molecular weight (MW) when analysed by SDS-polyacrylamide gel electrophoresis (PAGE)[4,5]. Figure 1 shows the patterns for rat T and B lymphocytes and thymocytes. T lymphocytes have bands of 170,000 and 95,000 apparent MW (Fig. 1, track 1) and B lymphocytes have one major band of 200,000 apparent MW (Fig. 1, track 2). Thymocytes have predominant bands of 150,000, 95,000 and 25,000 apparent MW (Fig. 1, track 3). The periodic acid Schiff (PAS) staining method identifies similar glycoprotein bands in lymphoid cell membranes analysed by SDS-PAGE[6].

Conventional or monoclonal antibodies against these glycoproteins were prepared and used for analysis and purification of the antigens. The 25,000 apparent MW glycoprotein of thymocytes is the Thy-1 antigen[7]. This molecule is an abundant glycoprotein of rodent thymus (10^6 molecules per thymocyte) and brain and occurs in two allotypic forms in the mouse (reviewed in ref. 8). The correct MW of thymocyte Thy-1 is 18,700 (ref. 9), of which 30% is carbohydrate. The polypeptide of rat brain Thy-1 has been sequenced and the attachment points for three *N*-linked carbohydrate structures have been identified[10]. The Thy-1 antigen seems to be structurally homologous to an immunoglobulin domain[11,12].

The 150,000, 170,000 and 200,000 apparent MW glycoproteins of thymocytes, T cells and B cells share antigenic determinants. These glycoproteins seem to be specific to haematopoietic cells and are synonymous with the leukocyte-common (L-C) antigen in the rat[13] and the T200 glycoprotein in the mouse[14]. In both rat and mouse these molecules bear allotypic determinants – the Ly-5 mouse alloantigen is the T200 glycoprotein[15] and the ART and Ly-1 rat alloantigens are the L-C antigen[16]. The rat thymocyte L-C antigen has been purified using a monoclonal antibody affinity column. There are about 70,000 molecules of antigen per thymocyte[13].

The 95,000 apparent MW bands found on rat thymocytes and T cells share an antigenic determinant recognized by a mouse anti-rat monoclonal antibody called W3/13 (ref. 5). This antibody was prepared after immunization of mice with rat thymocyte plasma membrane[17]. The W3/13 antigen is found on thymocytes, T (but not B) lymphocytes, bone marrow neutrophils and brain, and also on plasma and myeloma cells (ref. 17 and D. Secher, personal communication). The antigen from haematopoietic cells is in every case a glycoprotein of apparent MW 95,000 which is heavily labelled by NaB^{3}H$_4$ after mild periodate oxidation (ref. 5 and W.R.A.B., unpublished data for rat myeloma cells). Brain W3/13 antigen has not been characterized. The 95,000 MW glycoprotein is notable because it does not bind at all to lentil lectin[5,6] and also because removal of sialic acid with neuraminidase results in an apparent increase in size on SDS-PAGE. The asialo glycoprotein can be labelled with ^{3}H-borohydride after oxidation of galactose with galactose oxidase[4]. The results for rat thymocytes are shown in Fig. 1, track 6. The bands of apparent MW 150,000 and 25,000 are unaffected by neuraminidase treatment but the 95,000 apparent MW band is found at 105,000 apparent MW after removal of sialic acid (compare tracks 3 and 6 in Fig. 1). Both these forms of the molecule bind to affinity columns of W3/13 antibody attached to Sepharose 4B (Fig. 1, tracks 5 and 8). Note that not all of the 95,000 or 105,000 MW bands bind to the antibody column (Fig. 1, tracks 4 and 7); we will discuss this further below.

This paper reports the purification of the W3/13 (95,000 MW) glycoprotein and the amino acid and carbohydrate compositions of this molecule and also of the L-C antigen (glycoprotein 150,000 MW) from rat thymocytes. The compositions are compared with those previously published for rat thymocyte Thy-1 antigen (glycoprotein 25,000 MW)[18] and for glycophorin A, the major sialoglycoprotein of human erythrocytes[19].

---

* Present address: Nuffield Department of Obstetrics and Gynaecology, The John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK.

0028-0836/81/060456—05$01.00

Table 1 Purification of thymocyte W3/13 antigen

Fraction	Antigenic activity (units $\times10^{-4}$)	Protein (mg)	Relative specific activity	Yield of antigenic activity (%)
Cells (2.82×10^{11}) + Brij 96	9.7	4,825	1	100
Brij 96 extract	9.7	2,594	1.9	100
Brij 96-deoxycholate extract	9.0	1,979	2.3	93
Sepharose 4B–W3/13 antibody unbound fraction	0	1,802	0	0
Sepharose 4B–W3/13 antibody bound and eluted fraction (1st)	7.6	—	—	78
Sepharose 4B–W3/13 antibody bound and eluted fraction (2nd)	7.2	—	—	74
After Sephacryl S300 chromatography	3.7	—	—	38
After ethanol precipitation	2.8	0.198	7,000	29

Thymocytes from Sprague–Dawley rats (OLAC 76) were teased into Dulbecco's A and B buffer, *p*H 8.4. To 2.82×10^{11} cells in 170 ml were added 73 ml of 10% Brij 96 (Sigma), 10 mM iodoacetamide (IAA) and 10 mM diisopropyl fluorophosphate (DFP) in Dulbecco's A and B buffer, *p*H 8.4. The mixture was stirred on ice for 30 min, homogenized with six strokes of a Potter–Elvehjem homogenizer and then centrifuged at 6,300*g* for 20 min. The supernatant (205 ml, Brij 96 extract) was added to 102 ml of 3% sodium deoxycholate and 10 mM Tris-HCl, *p*H 8.0, and the mixture was centrifuged at 70,000*g* for 45 min. The supernatant (285 ml, Brij 96 sodium deoxycholate extract) was then passed at a flow rate of about 20 ml h^{-1} through a control column of mouse immunoglobulin coupled to Sepharose 4B and an 8-ml affinity column of 120 mg of the W3/13 antibody coupled to Sepharose 4B (refs 38, 39). After washing with 2% sodium deoxycholate, 10 mM Tris-HCl, 2 mM IAA and 2 mM DFP, the affinity column was eluted with 50 mM diethylamine–0.5% sodium deoxycholate, *p*H 11.5. The eluate was brought to *p*H 8.6 by the addition of glycine. This eluate was not pure W3/13 antigen and so was passed through the affinity column as above. The material specifically bound and eluted from the column a second time was then concentrated by ultrafiltration over an Amicon PM 10 membrane and chromatographed on Sephacryl S300 in 0.5% sodium deoxycholate and 10 mM Tris-HCl, *p*H 8.4. The eluted material containing the W3/13 antigen was concentrated before precipitation with 75% ethanol[18] at −20 °C. One unit of antigenic activity is the amount of antigen needed to give 50% inhibition of the indirect radioactive binding assay for W3/13 (ref. 17). Protein was assayed by the Lowry method[40] for detergent extracts and by amino acid analysis for the pure glycoprotein. All procedures were carried out at 4 °C.

Purification of the W3/13 antigen (glycoprotein 95,000 MW)

The W3/13 antigen was purified from rat thymocytes using a monoclonal antibody affinity column. In previous purifications of the Thy-1 (ref. 7) and L-C (ref. 13) antigens a plasma membrane fraction was first prepared and then solubilized in deoxycholate to give an extract which was passed through an antibody affinity column. Bound antigen was eluted from the column with high *p*H. This scheme posed problems with the W3/13 antigen because the glycoprotein was very susceptible to proteolysis, yielding a form with an apparent MW of 82,000 as shown in Fig. 1, track 9. This degradation occurred extensively when membranes were prepared using the Tween 40 method[6] even though neither Thy-1 nor L-C antigen showed autolysis during this procedure. The addition of iodoacetamide and diisopropyl fluorophosphate minimized this degradation (Fig. 1, track 10) and we also found that direct solubilization of membrane from whole cells with the non-ionic detergent Brij 96 gave less autolysis. This step also resulted in the recovery of more antigenic activity than when membranes were prepared. The only disadvantage of using the non-ionic detergent was that in antibody affinity chromatography more non-antigenic proteins were adsorbed to the columns than when deoxycholate was used. To minimize this, deoxycholate was added to the Brij 96 extract and the affinity column step was carried out twice.

Table 1 legend describes the purification of the W3/13 antigen. All the antigenic activity was bound by the column and 85% of the applied activity was recovered in the high *p*H eluate. After repeating the affinity column step and carrying out gel filtration, about 30% of the original antigenic activity was recovered at a final purification of ~7,000-fold. It is most unlikely that such a purification could have been achieved by any known procedure other than antibody affinity chromatography.

The purification factor is of the right order of magnitude for obtaining pure W3/13 antigen. If it is assumed that the whole

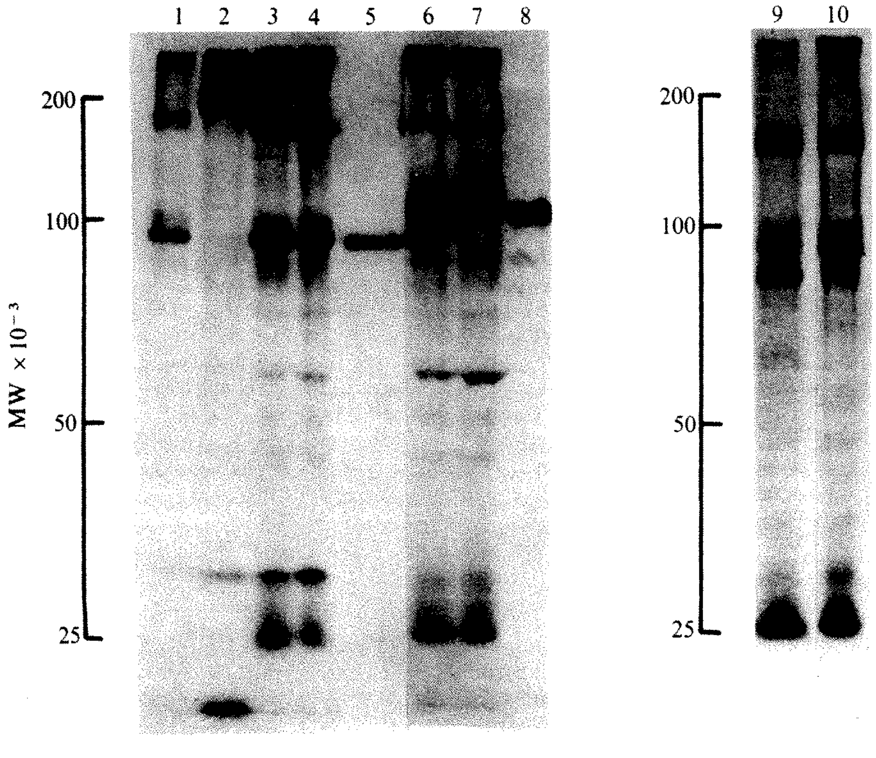


Fig. 1 SDS-PAGE of radiolabelled cell-surface glycoproteins. Tracks 1–5: glycoproteins labelled in sialic acid residues (see below). Track 1, from T lymphocytes; track 2, from B lymphocytes; track 3, from thymocytes; track 4, from thymocytes passed through a W3/13 antibody affinity column; track 5, W3/13 antigen eluted from affinity column. Tracks 6–8: glycoproteins labelled in galactose residues (see below). Track 6, from thymocytes; track 7, from thymocytes passed through a W3/13 antibody affinity column; track 8, W3/13 antigen eluted from affinity column. Track 9: thymocyte glycoproteins labelled in sialic acid from plasma membranes without inhibitors of proteolysis. Track 10: as for 9 plus inhibitors (see below). For tracks 1–8 thymocytes and T and B lymphocytes were prepared as in ref. 5 and labelled with NaB^{3}H$_4$ after oxidation with periodate[3] or galactose oxidase following neuraminidase digestion[4]. Extracts were prepared by resuspending the labelled cells at 3×10^8 ml^{-1} in 3% Brij 96 in Dulbecco's A and B buffer, 5 mM iodoacetamide and 5 mM DFP. After incubation for 15 min on ice nuclei were removed by centrifugation for 10 min at 2,000*g*. One vol of 3% sodium deoxycholate, 10 mM Tris-HCl, *p*H 8, was then added to 2 vol of the Brij 96 extract. For affinity chromatography the extract was passed through a column of W3/13 antibody and washed and eluted as in Table 1. The eluted material was brought to *p*H 8.6 with glycine and inhibitors of proteolysis were added. For tracks 9 and 10 plasma membrane was prepared from radiolabelled thymocytes using the Tween 40 method[6]. To minimize losses $\sim13.5\times10^8$ unlabelled thymocytes were added to 1.5×10^8 labelled thymocytes. For track 10 but not 9 inhibitors of proteolysis (as above) were added at each step. For all samples, glycoproteins were precipitated with 80% acetone at −20 °C overnight and the precipitate was recovered by centrifugation and suspended in 8 M urea, 5% SDS, 2% dithiothreitol. The samples were boiled for 5 min and run on 10% polyacrylamide gels[46]. The radioactive bands were visualized by fluorography[5].

molecule has a MW of 95,000, then that of the protein part is 38,000 (60% is carbohydrate, see below). From this, and the data in Table 1, it can be calculated that there should be 38,000 molecules of W3/13 antigen per thymocyte. At saturation, 35,000 molecules of the W3/13 monoclonal antibody bind per thymocyte[17] and this probably underestimates the number of antigenic sites by about 50% (ref. 20). Although both estimates of the amount of antigen suffer from some uncertainties, the data strongly suggest that the purified molecule is the W3/13 antigen.

The purified W3/13 antigen was analysed by SDS-PAGE followed by staining for protein with Coomassie blue (Fig. 2*a*) and carbohydrate with PAS stain (Fig. 2*b*). The purified glycoprotein stains very heavily with PAS stain and this is illustrated in Fig. 2 by comparing the staining of W3/13 antigen with that of rat brain Thy-1 antigen. With Coomassie blue stain 4 μg of protein gave roughly equal staining with both glycoproteins (Fig. 2*a*), whereas with PAS stain the W3/13 glycoprotein was stained much more intensely than Thy-1 (Fig. 2*b*). The purified W3/13 antigen showed more than one band with both Coomassie blue and PAS stain. Densitometry on the trace in Fig. 2*a* gave 70% of the stained material in the peak at 95,000 apparent MW, with other minor peaks at apparent MWs 86,000 and ~70,000. These are probably proteolytic products of W3/13 antigen as they stain heavily with PAS stain. This possibility was supported by an experiment in which the glycoprotein was run on SDS-PAGE followed by slicing the gel and assaying for W3/13 antigenic activity in material eluted from the gel. The results, which are considered to be semi-quantitative, are shown in Fig. 2*c* and antigenic activity is associated with the protein and carbohydrate bands shown in Fig. 2*a* and *b*. Small amounts of antigenic activity were also detected at an apparent MW of 58,000 where Coomassie blue staining bands cannot be seen. The overall conclusion from Fig. 2 is that the W3/13 antigen is sufficiently pure for determination of carbohydrate and amino acid compositions.

Table 2 Amino acid and carbohydrate composition of the W3/13, L-C and Thy-1 antigens and of human glycophorin A

Residue	W3/13	L-C	Thy-1	Glycophorin A
	Amino acids per 100 amino acid residues			
Asx	5.3	11.1	12.6	6.1
Glx	9.4	12.0	9.1	10.7
His	0.6	2.3	3.8	3.8
Lys	3.6	7.3	7.2	3.8
Arg	3.5	4.9	7.5	4.5
Thr	13.7	6.5	8.6	11.5
Ser	13.8	7.7	7.1	14.5
Pro	9.4	4.8	3.8	7.6
Ala	8.7	5.7	2.9	4.6
Cys	0	2.1	3.2	0
Gly	9.5	6.4	4.9	4.6
Tyr	0.5	3.8	2.0	3.1
Val	7.5	6.5	7.4	8.4
Ile	2.5	4.3	4.1	8.4
Leu	8.1	7.4	11.2	5.3
Phe	2.7	3.6	3.7	1.5
Met	1.4	1.6	0.9	1.5
Trp	0	2.1	0	0
	Carbohydrate residues per 100 amino acid residues			
Fucose	0	0.9	1.0	1.0
Mannose	0	4.4	10.6	4.0
Galactose	22.3	3.6	5.5	19.0
Glucose	11.1	1.1	1.3	0
Glucosamine	0	5.7	9.4	6.2
Galactosamine	20.1	1.1	0	11.1
Sialic acid	27.5	5.4	1.8	21.9
Sialic acid by fluorometric method	27.2	2.9	2.1	—
% by weight carbohydrate	60	25	32	56

The data for W3/13 and L-C antigens were determined as below; the data for rat thymocyte Thy-1 (L^+ form) and human glycophorin are from refs 18 and 19, respectively. (The glycophorin carbohydrate composition was calculated from that given for the chymotryptic 1 peptide, which contains all the carbohydrate structures.) Amino acids were quantified on a Durrum D500 after 24, 48 and 72 h of hydrolysis in 6 M HCl and 0.005 M phenol *in vacuo* at 110 °C. Ser and Thr were calculated by extrapolation to zero time. Cys and Met were determined after performic acid oxidation[41]. The data for W3/13 and L-C were for five and four preparations, respectively. Trp in L-C antigen was detected spectrophotometrically[42]. No Trp was detected in W3/13 by the method described in ref. 43. The s.e.m.s were less than 5% of the mean value except for Met, Tyr and His, which were less than 12% of the mean value. Hexoses and amino sugars were determined by gas–liquid chromatography after methanolysis and trimethylsilylation[44]. Sialic acid was also determined fluorometrically[45] with correction for interferences by hexoses and hexosamines. The carbohydrate analyses were from two to five separate preparations of W3/13 and L-C antigens. The s.e.m.s were less than 10% of the mean. The amino sugars were also analysed by ion exchange chromatography following hydrolysis in 3 M *p*-toluene sulphonic acid[43]. The values (not shown) were similar to those obtained by gas–liquid chromatography. Glucose was also found in fractions from the Sephacryl S300 column which did not contain W3/13 antigen and was therefore assumed to be a contaminant. The percentage by weight of carbohydrate of W3/13 and L-C glycoproteins was determined using the sialic acid values obtained by the fluorescence method, assuming it was *N*-acetylneuraminic acid. Amino sugars were assumed to be *N*-acetylated.

A final point regarding the purification is the relationship between the molecule which binds to the W3/13 antibody column and the whole of the 95,000 MW glycoprotein band. This glycoprotein appears homogeneous and the fact that all ^{3}H-labelled material shows the unusual behaviour after neuraminidase digestion (see Fig. 1) would suggest that the whole band is structurally similar. Despite this, it is clear that the W3/13 monoclonal antibody binds only a part of the glycoprotein. When solubilized membrane is passed through the affinity column, although all W3/13 antigenic activity is bound (Table 1) a part of the 95,000 apparent MW or the asialo 105,000 apparent MW glycoprotein band does not bind (Fig. 1, track 4 compared with 3, and track 7 compared with 6). Radioactivity counting in gel slices indicates that about 50% of the glycoproteins did not bind. This is not an artefact of glycoprotein radiolabelling because the same result was seen if material passing through the antibody column was analysed by PAS staining after SDS-PAGE (data not shown). It seems that the monoclonal antibody detects antigenic heterogeneity in glycoprotein molecules with an overall similarity of structure, a view supported by the fact that mouse antibodies raised against the purified W3/13 antigen can react with the 95,000 apparent MW band which has passed unretarded through the W3/13 antibody affinity column (W.R.A.B., unpublished observations).

Chemical composition of the W3/13 antigen and the L-C antigen from thymocytes

The pure W3/13 and L-C antigens, purified as previously described, were analysed for their carbohydrate and amino acid contents. The results are given in Table 2 where they are compared with the compositions of rat thymocyte Thy-1 antigen and human erythrocyte glycophorin A.

The W3/13 antigen has a carbohydrate composition quite unlike either of the other two heavily glycosylated molecules of the rat thymocyte membrane. Carbohydrate comprises 60% w/w of the molecule and consists solely of galactose, galactosamine and sialic acid. (Some glucose was also detected but this is probably due to contamination; see Table 2 legend.) No mannose or glucosamine was detected and thus the molecule lacks carbohydrate structures of the type which are *N*-linked to asparagine through *N*-acetylglucosamine and have a core of glucosamine and mannose residues[21]. The carbohydrate composition of W3/13 glycoprotein is what would be expected if the molecule contained structures of the type

$$\begin{tikzcd} \text{NANA} \arrow[r, "\alpha 2-3"] & \text{Gal} \arrow[r, "\beta 1-3"] & \text{GalNAc} \arrow[r] & \text{Ser/Thr} \\ & & \text{NANA} \arrow[u, "\alpha 2-6"'] & \end{tikzcd}$$

which are found attached to proteins by an *O*-glycosidic linkage to serine or threonine[21]. The absolute amount of galactosamine and galactose suggests that there are about 20 such structures per 100 amino acids and this is supported by the finding that Ser and Thr together constitute 27 mol % of the amino acids, an unusually high proportion.

The *O*-glycosidic linkage of carbohydrate to protein is alkali labpe whereas the *N*-glycosidic bond is not. Alkali effects a

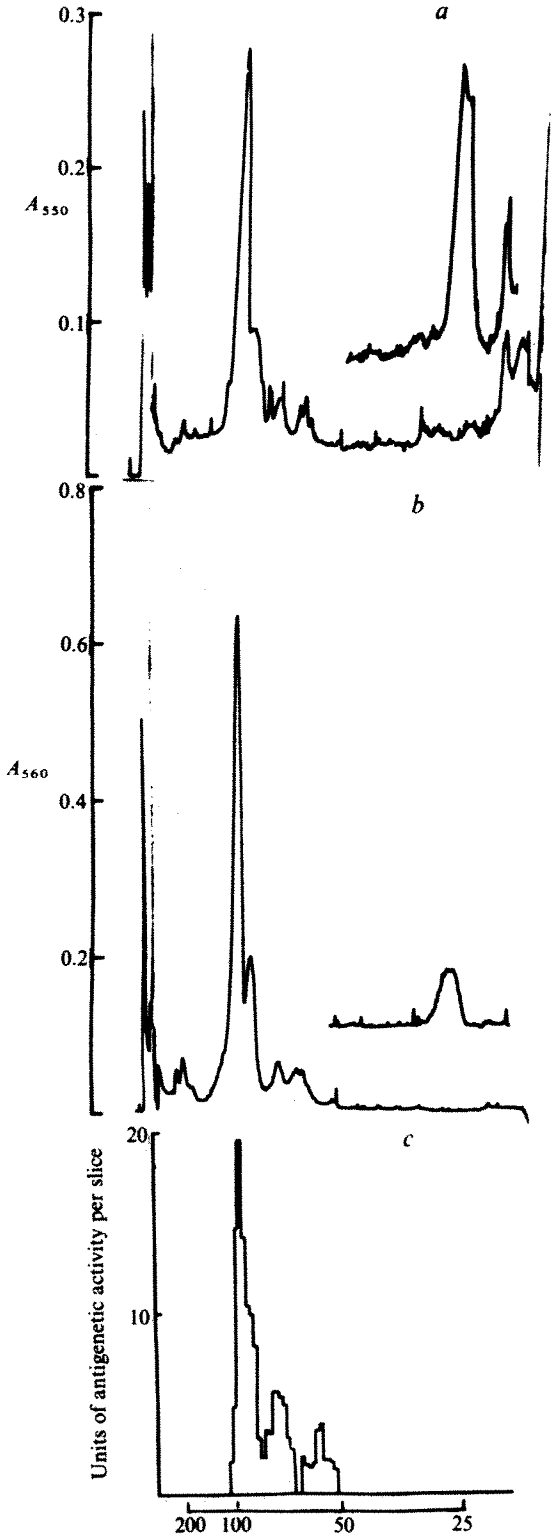


Fig. 2 SDS-PAGE of purified W3/13 antigen. W3/13 glycoprotein purified as in Table 1 was electrophoresed at 4 μg protein per sample on 11% polyacrylamide slab gels (3% stacking gel) and stained with Coomassie blue (*a*) and periodic acid Schiff stain (*b*) (methods as in ref. 6). The figure shows absorbance across the gel read by densitometry using a Gilford 2520 densitometer scanning at 550 nm for Coomassie blue and 560 nm for PAS stain. The insets show the traces from gels on which 4 μg protein of rat brain Thy-1 antigen was electrophoresed and stained with Coomassie blue (*a*) and PAS (*b*). *c* Shows the recovery of W3/13 antigenic activity after SDS-PAGE. A 3-μg protein sample of W3/13 glycoprotein was suspended in 0.1% SDS, 10% glycerol and electrophoresed as above. The gel was then sliced into 1-mm slices which were extracted overnight at 20 °C with 135 μl of 1% bovine serum albumin and 10 mM Tris-HCl, *p*H 8. The antigenic activity of the extracts was assayed by inhibition of an indirect binding assay for W3/13 antigen[17]. One unit of activity is the amount of antigen needed to give 50% inhibition of the standard assay. The total recovery of antigenic activity was 66% of that expected.

β-elimination of the carbohydrate attached to protein at serine or threonine residues. If this reaction is carried out in the presence of excess $NaBH_4$, the aldehyde at the reducing terminus of the eliminated oligosaccharide is reduced to an alcohol[22]. Thus, if the carbohydrate of W3/13 antigen is *O*-linked to Ser and Thr, it should be released in a low MW form by β-elimination in the presence of excess $NaBH_4$. This was examined using glycoprotein which had been labelled with tritium in its sialic acid or galactose residues (after removal of sialic acid). The results (Fig. 3) show that in both cases almost all the labelled carbohydrate was released by β-elimination into low MW forms which were retarded in gel filtration on a Bio-gel P-6 column. With asialo glycoprotein one low MW peak was obtained whereas alkali treatment of the sialylated glycoprotein gave two peaks. The latter two peaks presumably constitute structures containing one and two sialic residues. This is supported by the compositions (Table 2) which show that there is insufficient sialic acid for all structures to have two sialic acid residues as indicated in the structure above. The only unusual point about the results in Fig. 3 is that the ^3H-labelled oligosaccharides were retarded rather less than would be expected for their proposed MW in comparison with marker compounds chromatographed separately. The significance of this is unclear, but the data strongly suggest that all the sialic acid and galactose of the W3/13 antigen occurs in *O*-linked structures.

There is no similarity between the composition of the L-C antigen of rat thymocytes (glycoprotein MW 150,000) and W3/13 glycoprotein. Also, the amino acid content of L-C antigen shows no particular distinguishing features (Table 2); carbohydrate constitutes 25% w/w of the molecule, and the predominant saccharides present are mannose, galactose, *N*-acetylglucosamine and sialic acid. This suggests that the L-C antigen contains mainly *N*-linked carbohydrate, the low value for *N*-acetylgalactosamine arguing against a large number of *O*-linked structures. The compositions for L-C antigen are quite similar to those for Thy-1 antigen (Table 2). Brain Thy-1 antigen has three *N*-linked structures per polypeptide[10] (111 amino acids) and thymocyte Thy-1 may have the same attachment points but with structures differing in the extent of completion at the terminal branches[18]. The L-C antigen has about half as many mannose and *N*-acetylglucosamine residues per 100 amino acids as Thy-1 antigen. This suggests that there are roughly 1.5 *N*-linked carbohydrate structures per 100 amino acids, and thus about 15 *N*-linked structures per molecule, assuming a molecular weight of 150,000. The amino acid composition of Thy-1 antigen and L-C antigen are similar, but this is uninformative as the molecules differ markedly in size and do not have unusual amino acid compositions.

Discussion

We have shown here that one of the three major glycoproteins of rat thymocytes is unusual in having an *O*-linked carbohydrate structure attached to about one amino acid in five. This must have a major effect on the structure of the molecule. Presumably there would be little tertiary folding of the polypeptide and the molecule probably adopts an extended conformation in solution[23,24]. The expression of such a molecule at cell surfaces can be anticipated to have an important effect on the charge and topography of the outer surface. From the data in Tables 1 and 2 it can be calculated that on rat thymocytes 4×10^6 molecules of sialic acid are attached to the W3/13 antigen and probably twice this amount if all of the 95,000 MW glycoprotein is taken into account. This compares with $\sim 2 \times 10^6$ molecules of sialic acid per cell on Thy-1 antigen (two residues per molecule, 10^6 molecules per cell) and the same number on L-C antigen (30 residues per molecule and 70,000 molecules per cell). On both T lymphocytes and bone marrow neutrophils there is at least as much W3/13 antigen as on thymocytes[17]. An equivalent glycoprotein molecule can be identified on the cell surface of mouse thymocytes and T lymphocytes[3,4] and it seems possible that the major glycoprotein of 90,000 MW which has been identified on mouse neutrophils[25] is the equivalent of the rat W3/13 antigen.

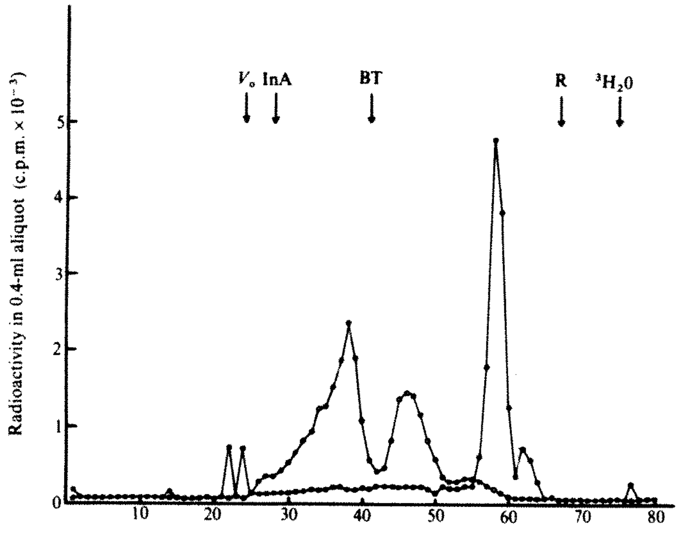


Fig. 3 Gel filtration of oligosaccharides derived by alkaline-borohydride degradation of ^{3}H-W3/13 antigen. ^{3}H-labelled W3/13 was prepared as described in Fig. 2 legend from thymocytes labelled with NaB^3H_4 after mild periodate oxidation (○) or after incubation with neuraminidase and galactose oxidase (●). The ^{3}H-labelled W3/13 was dialysed into distilled water and then subjected to alkaline borohydride degradation as described in ref. 22. After destruction of the excess $NaBH_4$ with acetic acid the mixture was fractionated on a Bio-gel P-6 column (102 × 1 cm) in 0.1 M NH_4HCO_3 at a flow rate of 4 ml h^{-1}. Fractions (1.1 ml) were collected of which 0.4-ml aliquots were scintillation counted in 3 ml Biofluor (NEN). The positions of elution of the following markers are shown. InA, insulin A, MW 2,360; BT, bacitracin, MW 1,460; R, raffinose, MW 504.

The expression of W3/13 antigen on T but not B lymphocytes may contribute to some of the differences between these cell types. Rodent T lymphocytes can be separated from B lymphocytes by electrophoresis because T cells are more acidic[26]. The presence of W3/13 antigen on T lymphocytes may be the reason for this difference in charge. Also, T lymphocytes are generally regarded as being less 'sticky' than B cells[27] and *in vivo* T cells recirculate more rapidly from blood to lymph[28]. Such differences could be mediated by an extended cell-surface molecule with a high concentration of negative charges which acts to impede surface contact.

Molecules with a high content of *O*-linked carbohydrate have been described from a variety of sources. A serum glycoprotein of this type is the anti-freeze protein of fishes which live at sub-zero temperatures[29]. A glycoprotein with similarities to the W3/13 antigen has also recently been purified from human serum[30]. Whether it might have originally derived from cell surfaces is not known. This molecule has an apparent MW of 140,000 on SDS-PAGE and contains 76% w/w carbohydrate, most of which is *O*-linked. Studies on this molecule suggest that it has a very extended conformation and little secondary structure[30]. Membrane or glycocalyx molecules with a high content of *O*-linked carbohydrates have also been described[21]. These include mucins which form a glycocalyx at epithelial cell surfaces and glycoproteins isolated from ascites tumour cells. These molecules seem to be much larger than the thymocyte glycoprotein described here and in the case of the tumour cell glycoproteins it has been suggested that the molecules mask other cell-surface structures[31,32].

The membrane molecule most like the W3/13 antigen is glycophorin, the major sialoglycoprotein of human erythrocytes. This molecule has 15 *O*-linked carbohydrate structures plus one *N*-linked structure attached to a sequence of 131 amino acids[19]. The amino acid and carbohydrate compositions for glycophorin are given in Table 2 and the similarities between this molecule and W3/13 glycoprotein are striking. Both molecules lack Cys and Trp and have low amounts of basic residues and high values for Ser, Thr and Pro. Significant differences are seen in the amounts of His, Ala, Gly, Tyr, Ile and Phe. In the carbohydrate compositions the *N*-linked structure of glycophorin is indicated by the presence of mannose and glucosamine. The similarities between the W3/13 glycoprotein and glycophorin suggest that these molecules may be evolutionarily related (alternatively, the similarities in amino acid composition could simply be a consequence of the unusual structure). There is more than one gene coding for the human erythrocyte glycophorins[33,34] and a family of genes, some members of which are expressed on other cells, could have evolved by gene duplication. The W3/13 antigen is not found on rat erythrocytes[17] but these cells are likely to express other glycophorin-like molecules[35].

The function of glycophorin is unknown and its importance has been questioned following identification of individuals who lack the molecule but show no defect in erythrocyte function[36,37]. This finding is difficult to evaluate because the expression of carbohydrate on other erythrocyte membrane molecules is altered in these individuals[36]. The finding of a glycophorin-like molecule as a major constituent of a number of other haematopoietic cell types (and possibly brain) strongly suggests that molecules of this type are functionally important. Their main role may be to alter the charge characteristics of the cell surface, particularly at a distance from the phospholipid bilayer. The expression of such molecules may affect the ease with which the membrane can align with that of another cell and thus be an important physical constraint in cell interactions.

We thank Mr W. Potts for assistance with carbohydrate analyses, Mr T. Gascoyne of the MRC Immunochemistry Unit, Oxford, for amino acid analysis and Mrs J. Kosak for typing the manuscript. W.R.A.B. was supported by an SRC research studentship.

Received 28 August; accepted 5 November 1980.

1. Hughes, R. C. *Essays Biochem.* **11,** 1–36 (1975).
2. Lloyd, C. W. *Biol. Rev.* **50,** 325–350 (1975).
3. Gahmberg, C. G. & Andersson, L. C. *J. biol. Chem.* **252,** 5888–5894 (1977).
4. Gahmberg, C. G., Häyry, P. & Andersson, L. C. *J. Cell Biol.* **68,** 642–653 (1976).
5. Standring, R., McMaster, W. R., Sunderland, C. A. & Williams, A. F. *Eur. J. Immun.* **8,** 832–839 (1978).
6. Standring, R. & Williams, A. F. *Biochim. biophys. Acta* **508,** 85–96 (1978).
7. Letarte-Muirhead, M., Barclay, A. N. & Williams, A. F. *Biochem. J.* **151,** 685–697 (1975).
8. Williams, A. F., Barclay, A. N., Letarte-Muirhead, M. & Morris, R. J. *Cold Spring Harb. Symp. quant. Biol.* **41,** 51–61 (1976).
9. Kuchel, P. W., Campbell, D. G., Barclay, A. N. & Williams, A. F. *Biochem. J.* **169,** 411–417 (1978).
10. Campbell, D. G., Gagnon, J., Reid, K. B. M. & Williams, A. F. *Biochem. J.* (in the press).
11. Campbell, D. G., Williams, A. F., Bayley, P. M. & Reid, K. B. M. *Nature* **282,** 341–342 (1979).
12. Cohen, F. E., Novotny, J., Sternberg, M. J. E., Campbell, D. G. & Williams, A. F. *Biochem. J.* (in the press).
13. Sunderland, C. A., McMaster, W. R. & Williams, A. F. *Eur. J. Immun.* **9,** 155–159 (1979).
14. Trowbridge, I. S. *J. exp. Med.* **148,** 313–323 (1978).
15. Omary, M. B., Trowbridge, I. S. & Scheid, M. P. *J. exp. Med.* **151,** 1311–1316 (1980).
16. Carter, P. B. & Sunderland, C. A. *Immunogenetics* **10,** 583–593 (1980).
17. Williams, A. F., Galfrè, G. & Milstein, C. *Cell* **12,** 663–673 (1977).
18. Barclay, A. N., Letarte-Muirhead, M., Williams, A. F. & Faulkes, R. *Nature* **263,** 563–567 (1976).
19. Tomita, M., Furthmayr, H. & Marchesi, V. T. *Biochemistry* **17,** 4756–4770 (1978).
20. Mason, D. W. & Williams, A. F. *Biochem. J.* **187,** 1–20 (1980).
21. Kornfeld, R. & Kornfeld, S. *A. Rev. Biochem.* **45,** 217–237 (1976).
22. Iyer, R. N. & Carlson, D. M. *Archs Biochem. Biophys.* **142,** 101–105 (1971).
23. Winzler, R. J. in *Red Cell Membrane, Structure and Function* (eds Jamieson, G. A. & Greenwalt, T. J.) 157–171 (Lippincott, Philadelphia, 1969).
24. Shulte, T. H. & Marchesi, V. T. *Biochemistry* **18,** 275–280 (1979).
25. Watt, S. M., Burgess, A. W. & Metcalf, D. *J. cell. Physiol.* **100,** 1–22 (1979).
26. Nordling, S., Andersson, L. C. & Häyry, P. *Science* **178,** 1001–1002 (1972).
27. Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3,** 645–649 (1973).
28. Howard, J. C. *J. exp. Med.* **135,** 185–199 (1972).
29. Komatsu, S. K., DeVries, A. L. & Feeney, R. E. *J. biol. Chem.* **245,** 2909–2913 (1970).
30. Schmid, K., Mao, S. K. Y., Kimura, A., Hayashi, S. & Binette, J. P. *J. biol. Chem.* **255,** 3221–3226 (1980).
31. Van den Eijnden, D. H. *et al. J. biol. Chem.* **254,** 12153–12159 (1979).
32. Sherblom, A. P., Buck, R. L. & Carraway, K. L. *J. biol. Chem.* **255,** 783–790 (1980).
33. Furthmayr, H. *Nature* **271,** 519–524 (1978).
34. Anstee, D. J., Mawby, W. J. & Tanner, M. J. A. *Biochem. J.* **183,** 193–203 (1979).
35. Sarris, A. H. & Palade, G. E. *J. biol. Chem.* **254,** 6724–6731 (1979).
36. Gahmberg, C. G., Myllyla, G., Leikola, J., Pirkola, A. & Nordling, S. *J. biol. Chem.* **251,** 6108–6116 (1976).
37. Tanner, M. J. A. & Anstee, J. D. *Biochem. J.* **153,** 271–277 (1976).
38. Axén, R. & Ernback, S. *Eur. J. Biochem.* **18,** 351–360 (1971).
39. Porath, J. *Meth. Enzym.* **34,** 13–29 (1974).
40. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193,** 265–275 (1951).
41. Hirs, C. H. W. *Meth. Enzym.* **11,** 59–62 (1967).
42. Edelhoch, H. *Biochemistry* **6,** 1948–1954 (1967).
43. Liu, T.-Y. & Chang, Y. H. *J. biol. Chem.* **246,** 2842–2848 (1971).
44. Clamp, J. R. *Biochem. Soc. Symp.* **40,** 3–16 (1974).
45. Hess, H. H. & Rolde, E. *J. biol. Chem.* **239,** 3215–3220 (1964).
46. Laemmli, U.K. *Nature* **227,** 680–685 (1970).

The reactivation of developmentally inert 5S genes in somatic nuclei injected into *Xenopus* oocytes

Laurence Jay Korn & J. B. Gurdon

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

In the frog Xenopus, *the somatic-type 5S ribosomal RNA genes are active in all cells; the oocyte-type 5S genes are active in oocytes but not in somatic cells. A new method of native gel electrophoresis resolves the two types of 5S RNA which are of the same length but different sequence. When somatic nuclei were injected into oocytes, their inactive oocyte-type 5S genes often remained inactive, and had thus conserved their developmentally regulated condition. Other conditions, which include treatment with 0.35 M NaCl, reactivated the genes.*

THERE are two classes of 5S ribosomal RNA genes in the genome of *Xenopus* and other amphibians. One class, comprising the oocyte-type 5S genes, is expressed in oocytes, but not in somatic cells[1,2]. The other class, the somatic-type 5S genes, is expressed in both oocytes and somatic cells[2]. We are investigating the mechanism by which the oocyte-type 5S genes are inactivated in somatic cells, as this may elucidate the mechanism by which the cell type-specific expression of genes is achieved in development.

The transcription of cloned 5S genes has been studied extensively with purified DNA in injected oocytes[3-5] as well as in cell-free systems[6-13]. However, in none of these cases is the natural pattern of differential transcription between oocyte- and somatic-type genes reproduced; in injected oocytes and somatic cell extracts, both types of 5S gene are transcribed. As we believe that genes in an intact nucleus are more likely to exhibit regulated transcription, we have injected whole somatic nuclei into oocytes. Then, if the oocyte-type 5S genes are activated, we can investigate changes associated with activation; but, if these genes remain inactive after injection, we can determine which components of the nuclei must be removed to obtain reactivation.

Previous work encourages this approach. We found that somatic nuclei injected into *Xenopus* oocytes are transcriptionally active for many days[14]. In one series of experiments, a gene product made in HeLa cells continued to be synthesized by HeLa nuclei injected into oocytes[15]. In other experiments, genes characteristically inactive in somatic cells were activated after the injection of nuclei into oocytes[16].

We describe here different conditions in which the oocyte-type 5S genes in somatic nuclei can either be made to remain inactive or can be reactivated after their injection into oocytes. To assay for selective transcription of oocyte- or somatic-type 5S RNA, we have developed a generally useful procedure for distinguishing between different RNAs of identical length. Although we do not yet know what factors are responsible for the developmental regulation of 5S gene transcription, this new approach to studying the expression of specific genes as part of an intact nucleus can be used to yield information on the control of these genes.

Rapid resolution of different RNAs of the same length

The oocyte and somatic types of *Xenopus* 5S RNAs are the same length (120 nucleotides) and differ from each other in sequence by six nucleotides in both *Xenopus laevis* and *Xenopus borealis*[17,18]. All four RNAs co-migrate on a denaturing polyacrylamide gel (Fig. 1*a*). The four types can be distinguished by RNA fingerprinting[17], but this procedure is too laborious for use with numerous samples. We have developed a simple electrophoresis assay which distinguishes all four 5S RNAs, and which is suitable for the analysis of many samples simultaneously.

X. laevis and *X. borealis* oocyte- and somatic-type 5S RNAs are unambiguously resolved by electrophoresis through a non-denaturing 17% polyacrylamide gel prepared and run as described in Fig. 1 legend. The substantially different mobilities of these four RNAs probably reflect differences in conformation in these conditions. In modified conditions which favour less secondary structure (elevated temperature, urea), resolution is considerably decreased. For our purposes, the main value of this method is for the rapid determination of the relative synthesis of oocyte- and somatic-types of 5S RNA (Fig. 1*b*).

The resolving power of this gel system, however, is not limited to 5S RNAs. The yeast tyrosine tRNA SUP4 and a mutant tRNA U67 are both 78 nucleotides long, and differ only by a single base change at position 67 ($C \rightarrow U$)[19]. These tRNAs migrate with different mobilities on native gels (Fig. 1*c*), though they migrate together on denaturing gels[20]. Similar results have been obtained for other tRNAs.

5S genes are efficiently transcribed in injected nuclei

Cultured cell nuclei injected into living oocytes are transcriptionally active. Oocytes injected with as few as 100–200 nuclei synthesize many times more somatic-type 5S RNA than uninjected control oocytes (Fig. 2, lanes 1–9 compared with lanes

Table 1 Activation of oocyte-type 5S genes in erythrocyte nuclei injected into oocytes

	Oocytes injected with erythrocyte nuclei	Adult somatic cells		
		Liver	Tissue culture cells	Factor of activation
Oocyte-type 5S RNA	40	<0.05	<0.07	800
Somatic-type 5S RNA	1	1	1	

Erythrocyte nuclei prepared by lysolecithin treatment were injected into oocytes, aiming for the germinal vesicle. Four hours later 1 μCi [α-^{32}P]GTP was injected, and the oocytes labelled for 2 days. Subsequent analysis was as for Figs 1, 2. The relative intensities of oocyte- and somatic-type 5S RNA bands were measured by microdensitometry. Data presented are the average of seven samples from two experiments. To determine the relative synthesis of oocyte- and somatic-type 5S RNA in somatic cells, an adult *X. laevis* was injected with 5 mCi $^{32}PO_4$. After 4 days, the liver was removed and the RNA extracted[36]. The sample was electrophoresed with a 5S RNA marker on a 12% polyacrylamide 7 M urea gel, the 5S RNA region excised and the RNA eluted from the gel[37]. The RNA was then re-electrophoresed on a 17% polyacrylamide native gel as described in Fig. 1 legend. *Xenopus laevis* kidney tissue culture cells were labelled for 24 h and the RNA isolated and electrophoresed as described in Fig. 3 legend. The relative intensities of oocyte- and somatic-type 5S RNA bands were determined. The absence of oocyte-type 5S RNA in tissue culture cells may be due to rapid turnover of such transcripts. However, we also failed to detect any oocyte-type 5S RNA when cultured cells were labelled for only 4 h.

0028-0836/81/060461—05$01.00

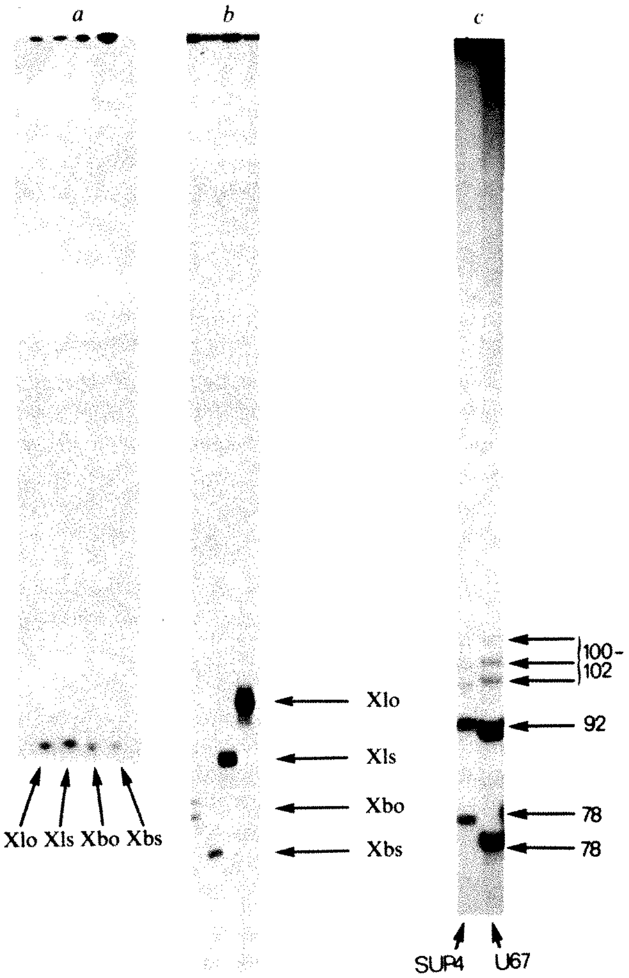


Fig. 1 Separation of RNA molecules by gel electrophoresis in non-denaturing conditions. *a*, Autoradiogram of a denaturing 12% polyacrylamide–7 M urea gel[38] of ^{32}P-RNA synthesized in *Xenopus* oocytes when [α-^{32}P]GTP and pXlo 8, pXls 11, pXbo 3, or pXbs 1 5S DNA were injected. pXlo 8 contains four *X. laevis* oocyte-type 5S genes[4]. pXbo 3 contains six *X. borealis* oocyte-type 5S genes[4,39]. pXls 11 and pXbs 1 contain one repeating unit each of *X. laevis* and *X. borealis* somatic-type 5S DNA, respectively[23]. About 5 ng DNA and 1 μCi [α-^{32}P]GTP were injected into the germinal vesicle of each oocyte. The oocytes were incubated at 20 °C for 24 h and the RNA extracted by the method of Brown and Littna[40]. The gel is 0.35 mm thick and contains 12% polyacrylamide, 7 M urea, 0.05 M Tris borate, *p*H 8.3, and 1 mM EDTA. For all gels, hypersensitized film was exposed at −70°C using a Dupont lightning plus intensifying screen. RNA transcripts from the cloned DNAs, following injection into oocytes, have been sequenced and identified as oocyte- or somatic-type 5S RNA from the respective *Xenopus* species[4,18,23]. *b*, Autoradiogram of a native 17% polyacrylamide gel of ^{32}P-RNA synthesized in *Xenopus* oocytes when [α-^{32}P]GTP and pXlo 8, pXls 11, pXbo 3, or pXbs 1 5S DNA were injected. DNA injections, and conditions for the synthesis and extraction of 5S RNA are as described for *a*. After ethanol precipitation and two 70% ethanol washes, the RNA sample was redissolved in 1–5 μl 10% glycerol, 1 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue. The 17% polyacrylamide (1:20 bisacrylamide/acrylamide), 0.35 mm-thick gel was prepared in 40 mM Tris, 20 mM Na acetate and 2 mM EDTA, to *p*H 7.8 with acetic acid. The electrode buffer was 50 mM Tris, 0.4 M glycine and 1 mM EDTA, *p*H 8.3. Electrophoresis was at constant current (9–11 mA) for 24 h. 5S RNA migrates through the gel at a rate of approximately 1 cm kV^{-1} h^{-1}. For maximum resolution, the gel should not be allowed to exceed 30°C. The pXbo 3 clone contains two types of Xbo gene (ref. 39, and L. J. K. and D. D. Brown, unpublished) the 5S RNA transcripts from these genes separate on the native gel. *c*, Autoradiogram of a native 17% polyacrylamide gel of ^{32}P-RNA synthesized in *Xenopus* oocytes when [α-^{32}P]GTP and pSUP4 or pU67 DNA were injected. pSUP4 contains the yeast Tyr tRNA SUP4 gene[41] and pU67 contains a single base change (C → U) at position 67 of the mature tRNA[19]. *Xenopus* oocytes transcribe and process these yeast tRNA genes[42]. The mature tRNAs are 78 nucleotides long. DNA injections, conditions for the synthesis and extraction of RNA and gel electrophoresis were as described above. tRNA precursor molecules of 92 and 100–102 nucleotides are indicated. Note that although the U67 mature tRNA migrates faster than SUP4, the 92- and 100–102-nucleotide precursor molecules do not. Therefore, the separation of the two 78-nucleotide molecules cannot be explained by the samples in one lane of the gel migrating faster than those in another.

10–12). The transcription of 5S genes in injected nuclei continues for at least 4 days (Fig. 2, lanes 7–9), and would probably continue for as long as the oocytes themselves remain synthetically active (sometimes for a few weeks). The rate of somatic-type 5S RNA synthesis has been estimated from the specific activity of the ^{32}P-GTP pool of oocytes (2.5 Ci mmol^{-1}), and from the number of ribosomes ($\sim 10^7$) synthesized per cell cycle in cultured cells. From these calculations, the rate of somatic-type 5S RNA synthesis in somatic nuclei injected into oocytes (500 transcripts per gene per h) seems to be at least as much as in the same nuclei when they were present in somatic cells (350 transcripts per gene per h). Thus this efficient transcription produces sufficient material from a single oocyte injected with nuclei to determine whether the 5S RNA synthesized is the oocyte or somatic type. As 5S transcription by injected nuclei is so much more than the endogenous level, the following results reflect the activity of 5S genes in the injected somatic nuclei and not that of the oocytes' own 5S genes.

The efficient transcription of nuclei injected into oocytes can be compared with that of isolated somatic nuclei incubated *in vitro*. Using two procedures optimized for transcription by RNA polymerase III (refs 6, 21), we found that 10^7 nuclei from *Xenopus* cultured cells incubated for 2 h synthesized only just enough 5S RNA to be detected (Fig. 2, TCN), or 0.025 transcripts per gene per h.

Activation or non-activation of oocyte-type 5S genes

Before assessing changes in 5S gene activity, we must establish the level of their activity before injection—in normal somatic cells and in oocytes. In somatic cells (see Table 1; Fig. 3, TCC; ref. 22) there are at least 20-fold more somatic-type 5S transcripts than oocyte-type 5S transcripts. This contrasts with the 5S gene copy number which in *X. laevis* favours the oocyte-type 5S genes by 50:1 (ref. 23). Thus, each somatic-type 5S gene is approximately 1,000 times more active in somatic cells than each oocyte-type gene. In the oocyte, on the other hand, the oocyte-type 5S transcripts predominate (oocyte/somatic 5S RNA ratio approximately 20:1) (ref. 2 and unpublished results). Note that, unlike the 18S and 28S ribosomal RNA genes, gene amplification is not responsible for this preferential expression of oocyte-type 5S genes in oocytes[24,25], the same number of these genes (20,000 per haploid chromosome set in *X. laevis*) being present in oocytes and in somatic cells.

As shown in Fig. 2, oocyte-type 5S genes of injected somatic nuclei can remain inactive after introduction into oocytes. Furthermore, this inactivity persists unaltered for at least 4 days. However, in several experiments we have observed the opposite result—an activation of oocyte-type 5S genes. Figure 3 shows the result of injecting samples of the same nuclear suspension into oocytes of two females. In the oocytes of one female,

Table 2 NaCl activation of oocyte-type 5S genes

	No. of individual assays			
	Lysolecithin		Triton	
Ratio of oocyte 5S/somatic 5S	−NaCl	+NaCl	−NaCl	+NaCl
>1 (activation)	14	7	1	14
<1 (no activation)	122	2	5	3
No. activated/no. not activated	0.11	3.5	0.2	4.7
Effect of NaCl $\left(\frac{\text{ratio}+\text{NaCl}}{\text{ratio}-\text{NaCl}}\right)$	31x		24x	

An activated oocyte is one in which, following injection of somatic nuclei, the amount of oocyte-type 5S transcripts is greater than somatic-type 5S. No activation means somatic-type 5S transcripts exceed oocyte-type ones. Nuclei were prepared from *X. laevis* tissue culture cells by lysolecithin or Triton treatment[26]. Salt-washed nuclei were incubated for 15 min at 0 °C in 0.35 M NaCl at a concentration of 5×10^6 nuclei ml^{-1} before injection. Approximately 500 nuclei were injected into single oocytes from non-activating females, aiming for the germinal vesicle. Each assay is a single oocyte.

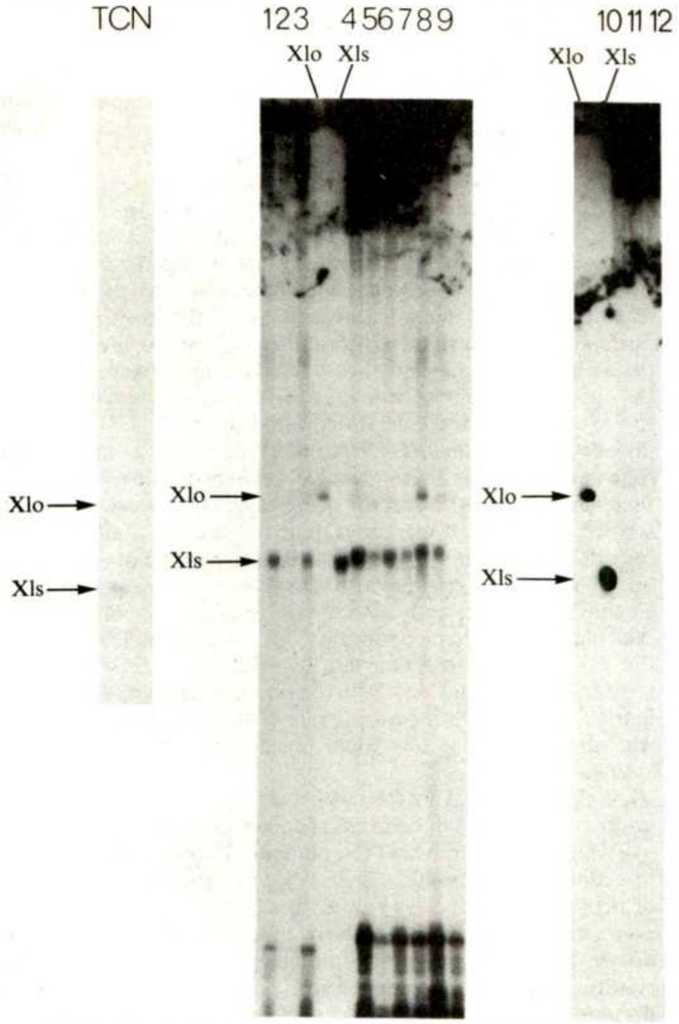


Fig. 2 5S gene transcription by somatic nuclei injected into oocytes. Tissue culture nuclei (TCN) were prepared by Triton treatment[26] from a line of *X. laevis* kidney cells growing in culture. 10^7 nuclei and 50 μCi [α-^{32}P]GTP were incubated for 2 h at 25°C as described by Crampton and Woodland[21]. The RNA was extracted and electrophoresed on a 12% polyacrylamide–7 M urea gel and the 5S RNA band was located by its co-migration with Xlo and Xls markers (see Fig. 1*a*). The 5S RNA band was excised and the RNA eluted[37]. The RNA was then electrophoresed on a 17% polyacrylamide native gel (track TCN). The positions of *X. laevis* oocyte-and somatic-type 5S RNAs which have been eluted from the denaturing gel and then electrophoresed on the native gel are indicated with arrows. Lanes 1–12, between 100 and 500 tissue culture cell nuclei prepared by lysolecithin treatment[26] were injected into individual oocytes, aiming for the germinal vesicle, followed by a second injection of 1 μCi [α-^{32}P]GTP at 4 h (lanes 1–3), 24 h (lanes 4–6) or 4 days (lanes 7–9) after the first injection. Tissue culture cells were removed from culture dishes by scraping rather than with trypsin because nuclei prepared from trypsin-treated cells were often toxic to oocytes. Control oocytes were injected with [α-^{32}P]GTP only (lanes 10–12). The endogenous oocyte transcription contributes to the high-molecular-weight RNA at the top of the gel. In this and all subsequent figures, the RNA was analysed as in Fig. 1 legend and half the RNA from individual oocytes electrophoresed in each lane of the native gel. tRNA molecules (bottom of gel) provide an internal control for the relative efficiency of labelling of RNA transcripts in different oocytes. Occasionally, a modest activation of oocyte-type 5S genes is observed (for example, lane 8); note, however, that Xlo 5S genes outnumber Xls genes by 50 : 1 (ref. 23).

Table 3 5S DNA-binding protein does not activate oocyte-type 5S genes

	Concentration (μg ml^{-1})	No. of oocytes with oocyte-type 5S genes Fully activated	Partially activated	Not activated
7S particle	1.5	0	0	7
7S particle	7.5	0	0	8
7S particle	60	0	2	5
Protein	0.8	0	0	8
Protein	3.8	0	0	8
Protein	30	0	0	8
None	—	0	1	7

Erythrocyte nuclei (about 200 per oocyte) were mixed with 7S particle or purified protein and injected into non-activating oocytes. The concentrations given are those of the material as injected. If the injected material was evenly dispersed throughout the oocyte, the intracellular concentration would have been about 10 times less (50 nl injected into 1 μl oocytes, of which about half the volume is solid yolk). The concentration of particle and protein used by Pelham and Brown[33] and in our experiments (purified as described in refs 32, 33) to give two- to six-fold stimulation of 5S DNA transcription ranged from <1 to >20 μg ml^{-1}. Injected ocytes are classified as partially activated when the ratio of labelled oocyte-type to labelled somatic-type 5S RNA was between 1 : 1 and 3 : 1, and as not activated when this ratio was 1 : 1 or less. Note that we never observed full activation (ratio $\geq$10 : 1). In a second experiment, in which again no evidence of activation was obtained by 7S particle or protein, full activation (oocyte-type/somatic-type 5S RNA ratio of $>$10 : 1) was observed when nuclei were treated with 0.35 M NaCl and Triton (as in Table 2 legend).

oocyte-type 5S genes were activated (Fig. 3, lanes 1–4); in the oocytes of the other they remained quiescent (lanes 5–8). Altogether we have tested oocytes from 50 females. Of these, 27 were activating, such that the amount of oocyte-type 5S RNA was between 5- and 40-fold greater than that of somatic-type 5S RNA. Sixteen females were non-activating, with an oocyte-type/somatic-type 5S RNA ratio of $<$1. Seven females had oocytes which gave intermediate values.

Note two aspects of these results. First, oocytes taken from the same female at intervals over a few months always (six females tested 2–4 times each) show the same activating or non-activating qualities as initially determined. Second, the pattern of somatic-type 5S transcription and tRNA synthesis in injected oocytes is approximately the same whether or not the oocyte-type 5S genes are activated (see, for example, Fig. 3). Somatic-type 5S RNA therefore provides an essential control, eliminating the possibility that the inactivity of oocyte-type 5S genes could be due to inviability of the injected nuclei in these oocytes. We conclude that the developmental inactivation of oocyte-type 5S genes takes place by a process that is readily reversed by components of the oocytes of some but not all females.

We have considered whether the oocyte-type 5S transcripts observed after the injection of somatic cell nuclei into activating oocytes could have resulted from a stimulation of transcription of the oocytes' endogenous 5S genes. This possibility has been eliminated by injecting *X. laevis* nuclei into oocytes from an activating *X. borealis* female. If the oocyte-type transcripts were coded for by the injected nuclei, they would be Xlo-type 5S RNA; however, if they were transcribed from the oocytes' own genes, they would be the Xbo type. The 5S RNA obtained co-migrates with Xlo, and not with Xbo (Fig. 3, lanes 9–11). Thus we conclude that the injected somatic nuclei are indeed activated to synthesize oocyte-type 5S RNA, and that the factor(s) responsible for this activation is not species specific.

5S gene activation in mature erythrocytes

Activation of 5S oocyte-type genes from cultured cell nuclei injected into oocytes might be due to the cell line becoming aberrant during many years of culture, as reported for a particular cultured line of *Xenopus* cells in which up to 15% of 5S transcripts are of the oocyte type[22]. We have therefore repeated our experiments with the nuclei of adult *Xenopus* erythrocytes to assess whether the inactivation of genes in a highly specialized, non-dividing, terminal cell type is also readily reversible.

Nuclei prepared by lysolecithin treatment[26] of mature erythrocytes of an adult frog were injected into oocytes of activating and non-activating females. Consistent with our results using nuclei from tissue culture cells, we find that erythrocyte nuclei injected into oocytes from non-activating females express somatic-, but not oocyte-type, 5S genes. In oocytes from activating females, on the other hand, the synthesis of oocyte-type 5S RNA was 40 times that of somatic 5S RNA (Table 1),

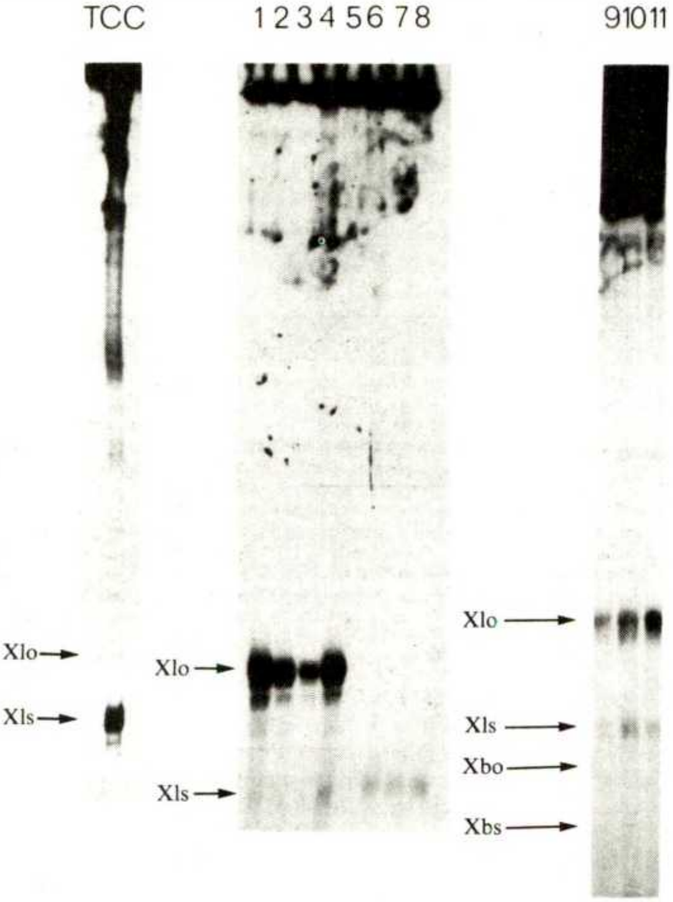


Fig. 3 Activating and non-activating oocytes. TCC, a 50% confluent flask of *X. laevis* kidney tissue culture cells (TCC) was incubated with 2 mCi ^{32}P inorganic phosphate for 24 h. Washed cells were lysed in isotonic buffer essentially as described by Kumar and Lindberg[43], except that the Nonidet P40 concentration was reduced to 0.2%. The RNA was electrophoresed on a 17% polyacrylamide native gel as described in Fig. 1 legend. Lanes 1–8, about 500 tissue culture cell nuclei prepared by lysolecithin treatment were injected with 1 μCi [α-^{32}P]GTP into individual oocytes from *X. laevis* female A (lanes 1–4) or female B (lanes 5–8), aiming for the germinal vesicle. Labelling was for 24 h. Lanes 9–11, approximately 300 *X. laevis* kidney TCC nuclei prepared by lysolecithin treatment were injected into individual *X. borealis* oocytes, aiming for the germinal vesicle, followed by a second injection of 1 μCi [α-^{32}P]GTP 4 h later. Labelling was for 40 h. The positions of *X. laevis* and *X. borealis* oocyte- and somatic-type 5S RNA, prepared as described in Fig. 1 legend, are indicated by arrows. tRNA synthesis, as seen at the bottom of the gel but not shown, was unaffected by the activated or non-activated state of the oocyte-type 5S genes.

just as was found in fully activated cultured cell nuclei. Similar results were obtained with a preparation of nuclei from immature blood cells taken from an adult *Xenopus* which had been rendered anaemic by the injection of phenylhydrazine 14 days earlier[27]. The 40-fold excess of oocyte over somatic 5S RNA represents a major activation of oocyte-type 5S genes, comparable with the endogenous ratio in oocytes (20–40 : 1). In somatic cells, as described above, the transcription of oocyte-type 5S RNA is typically less than 5% of that of the synthesis of somatic 5S RNA. We confirmed this by measuring the ratio of oocyte to somatic 5S RNA synthesis in liver, which is the erythropoietic organ of *Xenopus* adults (Table 1). Thus erythrocyte nuclei undergo at least an 800-fold activation of their oocyte-type 5S genes after injection into oocytes (see Table 1 legend for comments on this calculation). We conclude that the developmental inactivation of 5S genes is as easily reversed in a highly specialized, normal, adult cell type as it is in a cultured line of cells.

Oocyte components required for gene activation

The germinal vesicle is the site of 5S gene activation. To help identify the components responsible for inactivating oocyte-type 5S genes in somatic cells, we have investigated the intracellular location and general nature of the activity concerned. If the activity is localized in the oocyte's nucleus or germinal vesicle (GV), gene activation should occur only when nuclei are injected into the GV and not when injected into the cytoplasm. The site of injected nuclei has been determined by examination of stained oocyte sections, as well as by the co-injection of purified DNA known to be transcribed only in the GV[3,28]. In all the experiments reported above, injected nuclei were aimed at the oocyte's GV, and sections of oocytes revealed that in 90% of the cases, nuclei are indeed deposited there (Fig. 4). The co-injection of *X. borealis* DNA (Xbs) and whole nuclei was followed by analysis using native gel electrophoresis to distinguish Xb and Xl transcripts (Fig. 1*b*); Xbs DNA was transcribed in nearly all injections aimed at the GV, but never in injections aimed at the cytoplasm. The same type of analysis of nuclei injected together with DNA into activating oocytes established that the activation of oocyte-type 5S genes does not take place in the recipient oocytes' cytoplasm, though it did so in the GV of the same oocytes. Somatic-type 5S genes were transcribed by nuclei located in the cytoplasm as well as by those in the GV, although the overall rate of somatic 5S transcription was about 10 times less when nuclei were deposited in the cytoplasm.

Treatment with NaCl activates oocyte-type 5S genes. Another approach to the analysis of 5S gene activation by oocytes is to ask whether the removal of components of somatic nuclei before injection facilitates the effect. We have treated suspensions of nuclei in various ways before injecting them into oocytes of non-activating females, and have then assessed the extent of oocyte-type 5S activation by native gel electrophoresis. The results, which are summarized in Table 2, confirm that nuclei prepared by lysolecithin or Triton treatment alone only rarely synthesize more oocyte- than somatic-type 5S RNA. In contrast, 15 min incubation in 0.35 M NaCl at 0 °C, a procedure designed to remove most non-histone proteins[29], usually results in a pronounced activation of oocyte-type 5S genes (oocyte/somatic 5S ratio 10–30 : 1). In one experiment where nuclei were treated with 2 M NaCl and 0.5% Triton (Cook nucleoids)[30], a procedure which depletes almost all nuclear proteins, oocyte-type 5S genes were fully activated to give an oocyte/somatic 5S RNA ratio of over 30 : 1. Note that the 0.35 M NaCl treatment can activate the oocyte-type 5S genes to the same extent as in Cook nucleoids. Although the effect of removing materials from nuclei before injection is substantial, we have had occasional experiments in which NaCl treatment has little if any effect. We have not yet distinguished several possible explanations for the effect of NaCl.

The 5S DNA-binding protein does not cause activation. Engelke *et al.*[31] have isolated a protein from *Xenopus* ovaries that binds to the part of the coding region of 5S genes known as the control region[12,13], and which is required specifically for the transcription of both oocyte- and somatic-type 5S DNA *in vitro*. Recently, it has been shown that this same protein combines with 5S RNA to form a 7S particle characteristic of *Xenopus* ovaries[32,33]. These observations raise the possibility that this 5S DNA-binding protein may be responsible for 5S gene activation in oocytes.

We have tested this possibility by injecting oocytes of a non-activating female with somatic nuclei as well as with 5S DNA-binding protein or the 7S particle which contains this protein. We found no effect of either the protein or 7S particle on the reactivation of oocyte-type 5S genes in injected somatic nuclei (Table 3) or on total 5S transcription. We have tested the protein and particle on oocytes of two different non-reactivating females, and over several concentrations ranging from 1- to 25-fold the concentration required to give maximum stimula-

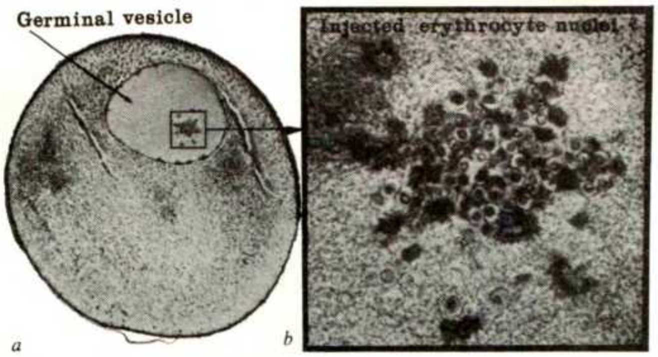


Fig. 4 Erythrocyte nuclei in injected oocytes. Oocytes were injected with a few hundred adult *Xenopus* erythrocyte nuclei prepared by the lysolecithin procedure[26], fixed after 1 day[26], sectioned and stained. *a*, The position of injected nuclei in the oocyte nucleus (germinal vesicle). *b* Shows an enlargement of the injected nuclei. Some of the nuclei appear to have formed dense-staining masses of chromatin, similar to those sometimes observed when HeLa nuclei are deposited in an oocyte germinal vesicle (see Fig. 3b of ref. 26). We do not know whether the 5S gene transcription which we observe comes from the chromatin masses or intact nuclei, both of which are seen in this oocyte section.

tion *in vitro*[33]. Pelham and Brown[33] showed that both protein and 7S particle independently stimulate transcription of 5S DNA *in vitro* by a factor of 2–6. As the preparations of protein and particle which we used also stimulated 5S DNA transcription when tested *in vitro*, we conclude that these substances are not sufficient for the reactivation of oocyte-type 5S genes in somatic nuclei.

Discussion

The rapid reactivation of the developmentally inert oocyte-type 5S genes in somatic nuclei, following their injection into oocytes, demonstrates the ready reversibility of gene expression in development, and is consistent with previous work involving the introduction of nuclei into *Xenopus* eggs or oocytes. The development of tadpoles from transplanted adult skin cell nuclei implies the reactivation of developmentally inert genes[34]. However, in this case, many cell divisions take place before reactivation is observed. More recently, two-dimensional gel analysis of proteins was used to demonstrate the synthesis of new proteins which are not expressed in a cultured line of *Xenopus* kidney cells, when nuclei from these cells were injected into oocytes[16]. The reactivation of gene expression is particularly clear in the case of the oocyte-type 5S genes, both because the primary transcripts derived from well characterized genes have been analysed, and the degree of developmental inactivation of these 5S genes is exceptionally strong. Nuclei injected into oocytes do not undergo division[26], and the activation therefore applies to the same actual genes or molecules of DNA as were previously inert. Our demonstration of 5S gene activation depends on the resolution of RNAs of identical size by gel electrophoresis in non-denaturing conditions, a method which is likely to be of considerable value in separating other groups of related RNAs such as low-molecular-weight nuclear RNAs[35].

The experiments reported here have shown that one can study the activity of genes in a regulated ('switched-off') condition. As far as we know, all cases in which purified genes as DNA have been tested for activity *in vitro* or *in vivo* result in the loss of the regulated state; for example both oocyte and somatic types of 5S DNA are fully active when injected into oocytes or added to somatic cell extracts[4,10,11]. This unique ability to observe regulated gene expression derives from the fact that in whole nuclei, as used in this study, genes are in their natural chromosomal environment. Thus the genes are still complexed with protein or other molecules with which they are naturally associated. The possibility that the developmental regulation of genes may require larger segments of DNA than are at present convenient to clone, or the cooperation of parts of the genome entirely unlinked to the genes being studied is easily accommodated by our system.

Our experiments give only indirect information about the nature of the activating effect of oocytes and hence about the developmental inactivation which is reversed. (1) Treatment with 0.35 M NaCl, which dissociates only a minority of chromosomal proteins[29], has nearly as much reactivating effect as 2 M NaCl and detergent which removes most of the molecules from DNA. This emphasizes again the lability of the developmental inactivation of 5S genes. We do not know whether NaCl activation involves the same mechanism as the effect of activating oocytes on untreated nuclei. (2) The 5S DNA-binding protein which promotes transcriptional activity *in vitro*[31,33] did not cause activation when tested by injection. Gene activation therefore seems unlikely to result from the mere presence of this protein in activating oocytes and its absence or low concentration in non-activating oocytes. This protein has not yet been shown to be able to discriminate between oocyte- and somatic-type 5S genes in its transcription-promoting effect.

The surprising existence of certain females which have non-activating oocytes, as well as others which have activating oocytes, was at first disturbing, but may eventually be providential; we can now prepare extracts of activating or non-activating oocytes and inject these with somatic nuclei into oocytes of the opposite type, with a view to fractionating the active extract.

We thank S. Whytock and R. Wells for assistance, Drs D. Brown, R. Roeder, K. Nishikura and E. De Robertis for communicating unpublished results, Dr J. Gottesfeld for providing *X. laevis* kidney tissue culture cells, and Drs D. Brown, R. Laskey, D. Melton, B. Meyer, G. Partington, J. Wells and M. Wickens for helpful discussions and comments on the manuscript. During part of this work, L.J.K. was a postdoctoral fellow of the Helen Hay Whitney Foundation.

Received 18 September; accepted 2 December 1980.

1. Wegnez, M., Monier, R. & Denis, H. *FEBS Lett.* **25,** 13–20 (1972).
2. Ford, P. J. & Southern, E. M. *Nature new Biol.* **241,** 7–12 (1973).
3. Brown, D. D. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2064–2068 (1977).
4. Brown, D. D. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2849–2853 (1978).
5. Gurdon, J. B. & Brown, D. D. *Devl Biol.* **67,** 346–356 (1978).
6. Birkenmeier, E. H., Brown, D. D. & Jordan, E. *Cell* **15,** 1077–1086 (1978).
7. Korn, L. J., Birkenmeier, E. H. & Brown, D. D. *Nucleic Acids Res.* **7,** 947–958 (1979).
8. Brown, D. D., Korn, L. J., Birkenmeier, E., Peterson, R. & Sakonju, S. *ICN-UCLA Symp.* **14,** 511–519 (1979).
9. Ng, S. Y., Parker, C. S. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **76,** 136–140 (1979).
10. Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18,** 469–484 (1979).
11. Weil, P. A., Segall, J., Harris, B., Ng, S. Y. & Roeder, R. G. *J. biol. Chem.* **254,** 6163–6173 (1979).
12. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. *Cell* **19,** 13–25 (1980).
13. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. *Cell* **19,** 27–35 (1980).
14. Gurdon, J. B., De Robertis, E. M. & Partington, G. A. *Nature* **260,** 116–120 (1976).
15. De Robertis, E. M., Partington, G. A., Longthorne, R. F. & Gurdon, J. B. *J. Embryol. exp. Morph.* **40,** 199–214 (1977).
16. De Robertis, E. M. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2470–2474 (1977).
17. Ford, P. J. & Brown, R. D. *Cell* **8,** 485–493 (1976).
18. Korn, L. J. & Brown, D. D. *Cell* **15,** 1145–1156 (1978).
19. Kurjan, J., Hall, B. D., Gillam, S. & Smith, M. *Cell* **20,** 701–709 (1980).
20. Nishikura, K., Kurjan, J., Hall, B. D. & De Robertis, E. M. (in preparation).
21. Crampton, J. M. & Woodland, H. R. *Devl Biol.* **70,** 453–466 (1079).
22. Ford, P. J. & Mathieson, T. *Nature* **261,** 433–435 (1976).
23. Peterson, R. C., Doering, J. L. & Brown, D. D. *Cell* **20,** 131–141 (1980).
24. Brown, D. D. & Weber, C. S. *J. molec. Biol.* **34,** 661–680 (1968).
25. Brown, D. D. & Dawid, I. B. *Science* **160,** 272–280 (1968).
26. Gurdon, J. B. *J. Embryol. exp. Morph.* **36,** 523–540 (1976).
27. Thomas, N. & Maclean, N. *J. Cell Sci.* **19,** 509–520 (1975).
28. Mertz, J. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **74,** 1502–1506 (1977).
29. Goodwin, G. H., Walker, J. M. & Johns, E. W. in *The Cell Nucleus* Vol. 6 (ed. Busch, H.) 181–219 (Academic, New York, 1978).
30. Cook, P. R., Brazell, I. A. & Jost, E. *J. Cell Sci.* **22,** 303–324 (1976).
31. Engelke, D. R., Ng, S. Y., Shastry, B. S. & Roeder, R. G. *Cell* **19,** 717–728 (1980).
32. Picard, B. & Wegnez, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 241–245 (1979).
33. Pelham, H. R. B. & Brown, D. D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4170–4174 (1980).
34. Gurdon, J. B., Laskey, R. A. & Reeves, O. R. *J. Embryol. exp. Morph.* **34,** 93–112 (1975).
35. Ro-Choi, T. S. & Busch, H. in *The Cell Nucleus* Vol. 3 (ed. Busch, H.) 151–208 (Academic, New York, 1974).
36. Shapiro, D. J., Baker, H. J. & Stitt, D. T. *J. biol. Chem.* **251,** 3105–3111 (1976).
37. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
38. Maniatis, T., Jeffrey, A. & Van Desande, H. *Biochemistry* **14,** 3787–3794 (1975).
39. Doering, J. Yb. *Carnegie Instn Wash.* **76,** 102–105 (1977).
40. Brown, D. D. & Littna, E. *J. molec. biol.* **20,** 95–112 (1966).
41. Goodman, H. M., Olson, M. V. & Hall, B. D. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5453–5457 (1977).
42. De Robertis, E. M. & Olson, M. V. *Nature* **278,** 137–143 (1979).
43. Kumar, A. & Lindberg, U. *Proc. natn. Acad. Sci. U.S.A.* **69,** 681–685 (1972).

Cruciform structures in supercoiled DNA

Nikos Panayotatos* & Robert D. Wells

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706

Specific sites in supercoiled pVH51 and pBR322 DNAs are cleaved by single-strand specific nucleases. These sites are in the loops of potential cruciform structures.

CRUCIFORMS have been favourite hypothetical recognition sites on DNA since they were first postulated[1]. The observation that several regulatory sites contain varying lengths of inverted repeat sequences encouraged these speculations[2]. If a cruciform existed *in vivo*, even transiently, it could be stabilized by a suitable protein. In supercoiled DNA, however, the necessary free energy could be provided by supercoiling and relatively stable cruciform structures might exist. The thermodynamic properties of superhelical DNA and potential cruciform structures have been studied[3–5].

This article describes strong evidence supporting two naturally occurring cruciform structures in supercoiled DNA.

Cleavage of supercoiled DNA by single-strand specific endonucleases

Figure 1 shows the rates of cleavage of supercoiled pVH51 DNA by S_1 nuclease[6] (*a*) and the T_7 gene 3 product[7–9] (*b*). In both cases, the supercoiled DNA (form I) is converted to the nicked-circular DNA (form II), which is then converted into the linear form (form III). Further nicking events on form III eventually cause double-strand breaks and generate smaller molecules which appear as a smear on the gel. Figure 1*a* shows that when relatively high levels of S_1 nuclease were used, complete conversion of the supercoiled DNA to form II was observed after only a few seconds. After 8 min, almost all of the nicked DNA had been linearized and by 90 min most had been degraded to smaller pieces. When lower levels of the T_7 gene 3 product were used, the gradual conversion of the supercoiled DNA to form II was observed over 4 min. Using appropriate levels of enzymatic activity, each of the discrete steps of the reaction can be monitored. Overall, pVH51 DNA is cleaved by S_1 nuclease and the T_7 gene 3 product with similar rates for each step.

Although these data permit a kinetic evaluation of the conversion of supercoiled DNA to linearized segments, they cannot provide information on the position-specificity of the cleavage, as both forms II and III are obligatory intermediates in the reaction.

Mapping of cleavage sites on pVH51 DNA

To determine if the nucleolytic cuts in supercoiled pVH51 DNA were at specific or random loci, the DNA was treated with each single-strand specific endonuclease and then specifically cleaved with a restriction enzyme. If the cleavage was the result of nicking at random sites, identical restriction patterns would be obtained for the single-strand specific nuclease-treated and untreated DNA. On the other hand, if linearization occurred at specific sites, the restriction pattern of the single-strand specific nuclease-treated DNA would show the loss of one (or more) fragment(s) and the appearance of new fragments with the same total length.

The results of these experiments are shown in Fig. 2. pVH51 DNA was incubated with S_1 nuclease until approximately one-half of the DNA was linearized (Fig. 2*a*, lane 2). (The reaction was not carried further so as to prevent the degradation that accompanies extensive linearization (see below).) When the S_1-treated and untreated DNAs were digested with *Hae*III, the gel analysis showed the typical pattern expected from this digest (ref. 10 and refs therein) (Fig. 2*b*). However, as shown in lane 2, the intensity of the 610-base pair band of the DNA that had been treated with S_1 nuclease was reduced by approximately 50% relative to the undigested control and, in addition, two new bands appeared. The sizes of the two new fragments were 340 and 270 base pairs and their sum (610 base pairs) equalled the size of the fragment whose intensity was reduced. This indicated that S_1 nuclease cleaves pVH51 DNA very specifically at a position inside the 610-base pair *Hae*III fragment.

When pVH51 DNA was linearized with the T_7 endonuclease and then restricted with *Hae*III, the same two new bands appeared at 340 and 270 base pairs (Fig. 2*c*, lane 5), indicating that this enzyme also cleaves pVH51 DNA at the same specific position as the S_1 nuclease. Furthermore, when analysed by a double digest with *Hae*III and *Hin*dII, *Hin*dII cleaved the 610-base pair fragment to the 510- and 100-base pair products expected from the known restriction map. The 340-base pair fragment was also cleaved to two new fragments 240- and

Table 1 Perfect inverted repeats

DNA	Base at centre of symmetry	No. of bases in stem	No. of bases in loop
pVH51	951.0	9	7
	1,009.0	9	5
	1,514.0	13	5
pBR322	3,220.5	10	6
	3,065.0	11	3
G4	2,671.5	8	15
	3,982.0	8	5
	2,571.5	9	0
ΦX174	3,008.5	8	8
	3,965.0	9	3
fd	4,698.5	8	12
	5,780.0	8	11
	5,803.0	8	3
	92.5	9	14
	1,548.5	9	4
	3,331.0	9	5
SV40	668.5	8	8
	2,026.0	8	7
	2,651.0	8	11
	5,092.0	8	4
	5,161.0	13	1

The DNA sequences of pVH51 (ref. 11 and J. Tomizawa, personal communication), pBR322[16], G4[28], ΦX174[29], fd[30] and SV40[20] were searched by computer for perfect inverted repeat sequences (only A+T and G+C pairs with no interruptions in the stems) which contain a maximum of 16 nucleotides in the loop regions. As expected, the number of inverted repeats decreases sharply as the stem length increases. Many perfect inverted repeats were found for all DNAs but only those with the two longest stem lengths are listed. Approximately 75% (2,694 base pairs) of the pVH51 sequence was available for this search.

---

* Present address: Biogen SA, Route de Troinex 3, 1227 Carouge, Geneva, Switzerland.

0028–0836/81/060466—05$01.00

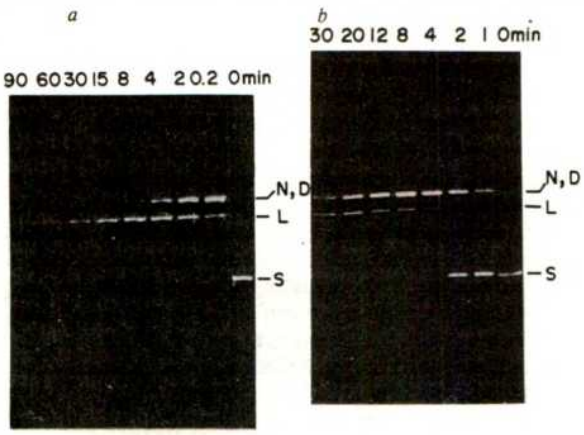


Fig. 1 Effect of single-strand specific endonucleases on supercoiled DNA. *a*, S_1 endonuclease; *b*, T_7 endonuclease. Reactions with S_1 nuclease were carried out in 40 mM sodium acetate buffer, *p*H 4.6, 0.05 M NaCl and 1.0 mM $ZnSO_4$. pVH51 DNA (5 μg) was reacted with 8 U of S_1 nuclease (specific activity 120,000 U per mg protein) in a final volume of 0.3 ml. The characterization of this preparation of S_1 nuclease has been described[21]. Reactions with T_7 endonuclease were carried out in a buffer containing 50 mM Tris-HCl, *p*H 7.6, 0.1 mM dithiothreitol, 10 mM $MgCl_2$ and 2.5 mM spermidine (hereafter called T_7 buffer). pVH51 (4 μg) was reacted with 6 U enzyme (specific activity 12,000 U per mg protein) in a final volume of 0.3 ml. T_7 endonuclease is the product of bacteriophage T_7 gene 3 (refs 7–9) and was purified from T_7-infected *E. coli* cells. The final preparation exhibited one major band, constituting at least 80% of the protein, on SDS-polyacrylamide gels (N.P. and R. D.W., unpublished results). One unit of activity is defined as the amount of enzyme necessary to convert 1 μg of supercoiled DNA into a mixture of nicked-circular and linear forms in 10 min in the conditions described above. All enzymatic reactions were carried out at 37 °C. Aliquots (30 μl) were removed, made 10 mM in EDTA at the indicated times and the mixtures were electrophoresed on 1% agarose gels and stained as described previously[22]. The positions of the supercoiled (S), linear (L), nicked (N) and dimer (D) DNA molecules are indicated. pVH51 DNA[23] was prepared as described previously[22].

100-base pairs long (Fig. 2*c*, lane 2). Thus, the unique linearization site must be located 240 ± 4 base pairs from the single *Hin*dII site of pVH51 DNA, which is inside the *Hae*III 610-base pair fragment.

The presence of a specific linearization site on pVH51 DNA was demonstrated with a nuclease-treated DNA preparation that contained approximately half form II and half linear molecules. At that ratio, nonspecific DNA degradation is minimal. However, if the nuclease reaction was allowed to proceed further to the point where all form II molecules were linearized, roughly 50% of the DNA was degraded to linear molecules of smaller than full-length size (Fig. 1*a*). On restriction of such a reaction mixture, the degraded molecules appeared on polyacrylamide gels as a background of numerous faint bands or even as a smear, depending on the progress of the reaction (data not shown). This indicates that, in addition to the highly specific linearization reaction, these single-strand specific nucleases are capable of degrading duplex DNA at a lower rate, by cleavage at less specific sites.

A 31-base pair inverted repeat sequence spans the linearization site

The discrete sizes of the 340-, 270- and 240-base pair fragments (Fig. 2) suggest that linearization results from nicks on the complementary strands at rather precise locations. If cleavage actually occurred at unique positions, it should be possible to end-label both strands at the cleavage sites and determine their nucleotide sequence. On the other hand, if linearization exposed more than one nucleotide at each end, labelling would not be specific and the sequencing gel would be impossible to read.

pVH51 DNA was linearized with S_1 nuclease, end-labelled and sequenced as described in Fig. 3 legend. Clear patterns were observed on the sequencing gel indicating that cleavage occurred at unique positions on each strand. Comparison of these sequences with the published sequence of pVH51 DNA (ref. 11, and J. Tomizawa, personal communication) revealed that the unique ends generated by S_1 nuclease were located exactly opposite each other on the two strands. Furthermore, the cleavage site was located 238 base pairs from the single *Hin*dII site of pVH51, in full agreement with the mapping data obtained with both the S_1 and T_7 nucleases.

Figure 3 shows the DNA sequence around the site of linearization. The remarkable feature of this sequence is an unusually long (13 base pairs) and perfect (no interruptions or G · T pairs) inverted repeat. The sequence can be shown in the form of a cruciform structure having double-stranded stems that are 13 base pairs long and single-stranded loops consisting of five nucleotides. The nicking site on each strand is located in the middle of each loop.

From these data, the following mechanism is proposed for the position-specific linearization of pVH51 DNA by the single-strand specific nucleases. The enzymes recognize a unique structural feature and cleave one strand, converting the supercoiled molecule into form II. Subsequently, the enzymes cleave the complementary strand exactly opposite the nick at a slower rate. At present, it is uncertain whether one of the complementary strands is cleaved first preferentially. Studies on the nonspecific nicking of PM2 DNA and ΦX174 RFI DNA by the

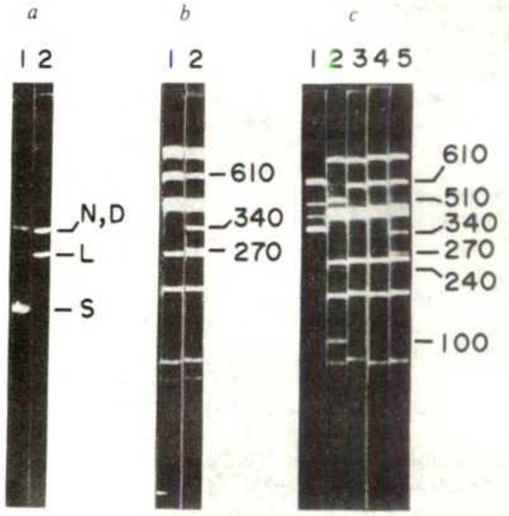


Fig. 2 Mapping of the cleavage site. *a*, Linearization of supercoiled pVH51 DNA with S_1 nuclease. Lane 1, 0.2 μg pVH51 DNA; lane 2, 0.4 μg pVH51 DNA reacted with S_1 nuclease. pVH51 DNA (60 μg) in a final volume of 1.0 ml was reacted with S_1 nuclease (15 U) at 37 °C for 5 min. The reaction mixture was extracted with phenol, followed by ether, and dialysed against 10 mM Tris-HCl, *p*H 8.0, 0.05 mM EDTA overnight. Aliquots of this mixture were used for subsequent reactions with restriction enzymes. Samples were analysed on 1% agarose gels. *b*, Mapping the S_1 cleavage site. Lane 1, 2 μg pVH51 DNA digested with *Hae*III; lane 2, 2 μg pVH51 digested with S_1 nuclease followed by *Hae*III. *c*, Mapping the T_7 endonuclease site. Lane 3, pVH51 DNA (2 μg) reacted with *Hae*III for 3 h; lane 4, pVH51 reacted with *Hae*III for 3 h followed by T_7 nuclease for 30 min; lane 5, pVH51 reacted with T_7 nuclease for 30 min, followed by *Hae*III for 3 h; lane 2, pVH51 reacted with T_7 nuclease for 30 min, followed by *Hae*III and *Hin*dII for 3 h; lane 1, a mixture (1 μg) of restriction fragments (365–610 base pairs in length; isolated for another purpose from pVH51 DNA by RPC-5 column chromatography[24]) incubated with T_7 nuclease for 3 h. All reactions were carried out in 40 μl at 37 °C in T_7 buffer. The second enzyme was added in a small (2 μl) volume. The sources and reaction conditions for all restriction enzymes used have been described[10,22,24–26]. The analyses in *b* and *c* were on 5% polyacrylamide gels. The sizes[10] (in base pairs) of selected restriction fragments are indicated.

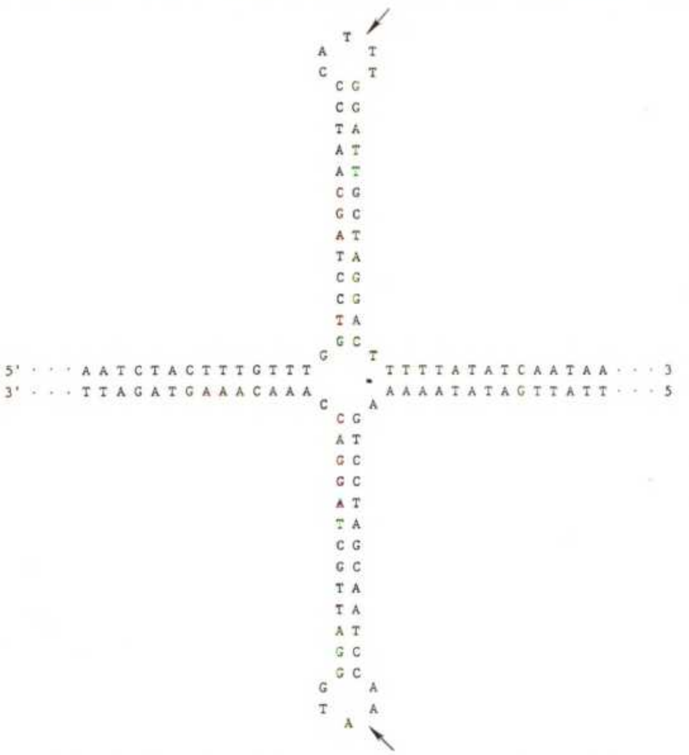

5' · · · A A T C T A C T T T G T T T
3' · · · T T A G A T G A A A C A A A
T T T T A T A T C A A T A A · · · 3
A A A A T A T A G T T A T T · · · 5

Fig. 3 pVH51 DNA sequence around the site cleaved by the single-strand specific endonucleases, drawn in cruciform structure. The positions of cleavage for each strand are indicated by arrows. pVH51 DNA was linearized with S_1 nuclease as described in Fig. 2 legend. Following treatment with bacterial alkaline phosphatase, the DNA was labelled with [γ-^{32}P]ATP, restricted with *Hae*III and fractionated on 5% polyacrylamide gels. Autoradiography revealed that 80–90% of the radioactivity was associated with the 340- and 270-base pair fragments. These fragments were eluted from the gel and sequenced as described[27] with some minor modifications[25]. Segments of both strands were sequenced and were in complete agreement with the published sequence[11].

Neurospora nuclease have shown that the first nick is not strand specific[5,12]. Likewise, it is uncertain whether one enzyme molecule cleaves both strands or whether the enzyme dissociates from the substrate after the first nicking event.

As described above, mapping studies on the site of cleavage with the T_7 gene 3 product demonstrated that both this enzyme and S_1 nuclease cleave at an identical (±4 base pairs) site. As the T_7 endonuclease was assayed at *p*H 7.6 in the presence of 10 mM $MgCl_2$, the possibility that the position-specific cleavage is induced by the unusual environment of the S_1 reactions (*p*H 4.6, 1 mM $ZnSO_4$) can be ruled out.

Supercoiling is necessary for site-specific cleavage

To determine the effect of supercoiling on the linearization rate, pVH51 DNA containing varying numbers of negative, superhelical turns was treated with the T_7 nuclease. The supercoil density and the appearance of linear molecules were monitored on agarose gels. Figure 4 shows that topoisomers within a higher density of supercoil turns are better substrates than the partially relaxed molecules. As the population of substrate DNAs with less supercoil turns increases (Fig. 4*b*), the amount of the linear molecules which are formed during the cleavage reaction decreases (Fig. 4*a*). In fact, the decrease in intensity of the band corresponding to linear molecules (Fig. 4*a*) closely parallels the decrease in intensity of the band of the most supercoiled molecules (Fig. 4*b*). In addition, partially relaxed molecules were much more resistant to the T_7 nuclease (Fig. 4*a*). To verify that the linear molecules observed in Fig. 4*a* resulted from site-specific cleavage, the DNA was restricted with *Hae*III and fractionated on polyacrylamide gels (Fig. 4*c*). The 340- and 270-base pair fragments were observed, as expected, and their intensities paralleled the intensity of the linear molecules observed in Fig. 4*a*.

Likewise, DNA restriction fragments, containing the sequence which is specifically cleaved in supercoiled DNA, are not cleaved in parallel conditions. Figure 2*c*, lane 4, shows that when pVH51 DNA was first restricted with *Hae*III and then treated with the T_7 nuclease, the 340- and 270-base pair fragments were not produced from the 610-base pair fragment. The same lack of specific cleavage was also observed when a mixture of restriction fragments including the 610-base pair fragment was incubated with the T_7 enzyme for prolonged times (Fig. 2*c*, lane 1). Hence, the DNA must be in a supercoiled, not relaxed or linear, conformation to be cleaved specifically by these single-strand specific nucleases.

Previous studies with the *Neurospora* endonuclease[5], mung bean nuclease[5], S_1 nuclease[13] and *Alteromonas espejiana* BAL 31 nuclease[14] have also shown that supercoiled DNA is a preferred substrate for the initial nick (conversion of form I to II). In contrast, the rate of nicking of PM2 DNA by pancreatic DNase I is independent of the degree of supercoiling[5].

Although a remarkably long, inverted repeat sequence exists around the cleavage site and the cuts are in the loops of the potential cruciforms, these data do not directly prove the presence of a cruciform structure. Two possible alternative recognition features (modified nucleotides and R-loops) have been considered and ruled out.

The presence of some types of modified nucleotides in pVH51 DNA specifically at the cleavage site could cause a partially disrupted DNA helix. To test this hypothesis, pVH51 DNA was grown in two different *Escherichia coli* host cells[10], M0 and

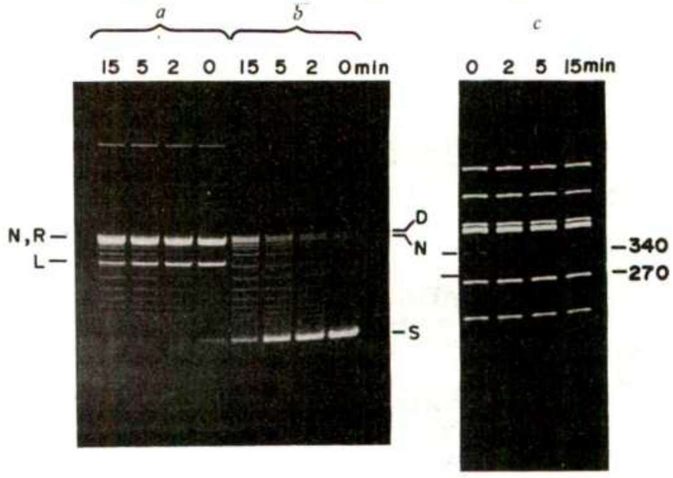


Fig. 4 Effect of supercoiling on the cleavage reaction. *b*, pVH51 DNA incubated with a nicking-closing enzyme for the indicated times; *a*, the partially relaxed DNA shown in *b* after incubation with T_7 endonuclease for 3 min; *c*, the partially relaxed and cleaved DNA shown in *a* after restriction with *Hae*III. pVH51 DNA (22 μg) was reacted at 37 °C with 4 μg wheat-germ topoisomerase in 60 μl T_7 buffer containing 30 mM NaCl. At the indicated times, each reaction mixture was placed at 72 °C for 10 min to inactive the enzyme and the solution was made 0.1 M with NaCl. Aliquots (10 μl) were then removed for electrophoresis on 1% agarose (*b*). T_7 endonuclease (1.2 U) was then added and the mixtures incubated at 45 °C for 3 min, followed by 10 min at 72 °C to inactivate the enzyme. Aliquots (10 μl) were removed for electrophoresis on 1% agarose gels (*a*). *Hae*III (2 U) was then added to each mixture and incubated for 3 h at 37 °C before electrophoresis on 5% polyacrylamide gels (*c*). The latter designations are as in Fig. 1 and R indicates the positions of relaxed, covalently closed, circular DNA. The wheat-germ DNA topoisomerase I was isolated and purified to approximately 60% purity by W. Dynan, J. Jendrisak, D. Hager and R. Burgess (unpublished method; gift of R. Burgess). This enzyme is free of nuclease activity and the products of a complete reaction are covalently closed, circular DNA without detectable amounts of nicked molecules. The faint bands on the top portion of *a* correspond to the form II-dimer and linear-dimer forms.

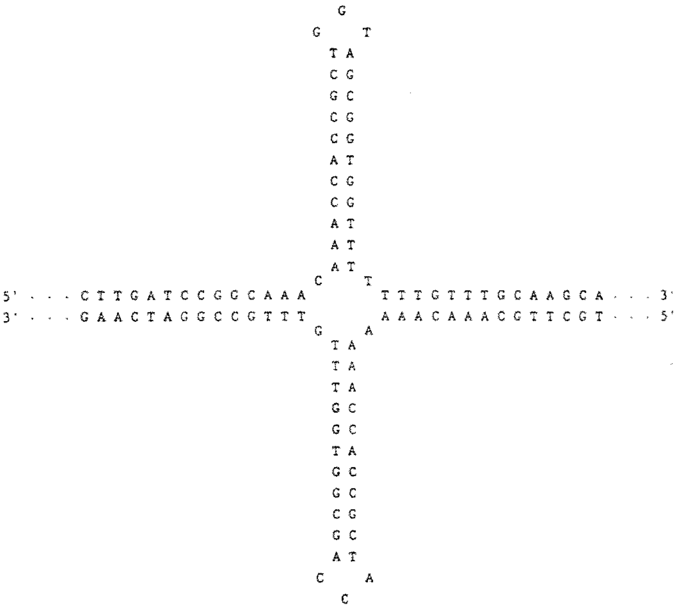

G
G T
T A
C G
G C
C G
C G
A T
C G
C G
A T
A T
A T
C T
5' . . . C T T G A T C C G G C A A A
T T T G T T T G C A A G C A . . . 3'
3' . . . G A A C T A G G C C G T T T
A A A C A A A C G T T C G T . . . 5'
G A
T A
T A
T A
G C
G C
T A
G C
G C
C G
G C
A T
C A
C

Fig. 5 pBR322 DNA sequence around the site cleaved by the single-strand specific endonucleases, drawn in a cruciform structure. The sequence is taken from ref. 16.

C600 SF8, the latter strain lacking a known modification system. Identical specific cleavages were found for both DNAs.

The second possibility is that an R-loop[15] might be present at the linearization site in the supercoiled DNA. This hypothesis was also eliminated when heat-denatured pVH51 DNA (which loses its R-loops[15]) produced the same results as native DNA.

Site-specific cleavage of supercoiled pBR322 DNA

If specific cleavage sites existed on supercoiled molecules other than pVH51 DNA, a comparison of their sequences might provide evidence as to the nature of the structural features necessary for recognition.

Supercoiled pBR322 DNA was linearized with the T_7 endonuclease and subsequently restricted in the manner described for pVH51. Two new bands appeared in the restriction patterns of the T_7 nuclease-treated DNA, indicating the presence of a specific cleavage site. The use of several restriction enzymes (*Hin*dII, *Hin*dIII, *Taq*I, *Hin*fI and *Hae*II) allowed rather precise mapping of the cleavage site at position 3,061 ± 5 base pairs of the published sequence of pBR322[16]. A *Taq*I digest of pBR322 DNA that had been treated with S_1 nuclease produced the same result.

Inspection of the DNA sequence revealed the presence of an inverted repeat around the cleavage site with a centre of symmetry at position 3,065. This repeat could be drawn into a cruciform configuration having double-stranded stems that are 11 base pairs long, and single-stranded loops consisting of three nucleotides (Fig. 5). The position of the loops corresponds to the position of cleavage, within experimental error. Therefore, as with pVH51, site-specific cleavage of pBR322 occurs at a position surrounded by an unusually long, inverted repeat sequence. Comparison of the DNA sequences around the cleavage sites in pVH51 and pBR322 revealed no significant homology that might constitute potential recognition sequences. Moreover, the requirement for superhelicity and the fact that two unrelated enzymes with no known sequence specificity cleave at the same site have already ruled out the possibility that the DNA sequence is recognized as a double-stranded helix. Instead, these results strongly support the notion that the supercoil pressure induces a tertiary (cruciform) structure which is maintained long enough for the single-strand specific nucleases to cleave at the single-stranded loops.

Absence of site-specific cleavage of other DNAs

Several other supercoiled DNAs were screened for position-specific cleavage. The replicative forms of G4, ΦX174, fd and M13 DNAs, as well as SV40 DNA, were linearized with the T_7 nuclease; S_1 nuclease linearization on SV40 and ΦX174 RFI DNAs was also studied. The nicking and linearization rates were comparable with the rates observed with pVH51 DNA (data not shown). Each DNA was then cleaved with at least two different restriction enzymes, in the manner described for pVH51 DNA, to ensure that the potential linearization site(s) did not overlap with a restriction site. In some cases, several bands were observed, indicating simultaneous linearization at many sites, but no unique bands that would allow unequivocal mapping of a cleavage site were found. Similar results have been published for the cleavage of ΦX174 RFI DNA by the *Neurospora* nuclease[12]. Therefore, G4, ΦX174, fd and M13 DNAs do not possess the unique structural features which, in the same conditions, lead to site-specific cleavage of pVH51 and pBR322 DNAs.

As the linearization sites on both pVH51 and pBR322 are in the middle of unusually long, inverted repeats, the lack of site-specific cleavage of G4, ΦX174, fd and M13 DNAs might be due to the absence of long, inverted repeats from the sequences of these molecules.

The published sequences of all six DNAs were searched by computer for inverted repeats and for A+T-rich regions.

Table 1 shows that pVH51 DNA contains the longest, perfect inverted repeat in the total of approximately 30,000 base pairs; this potential cruciform (Fig. 3) is the site of specific cleavage. The second longest repeat (13-base pair stem, one-base loop) was found in SV40 DNA. However, considering that a loop of less than three to four bases is sterically constrained[17], the maximum stem size for a cruciform of this repeat is 11 base pairs. The next longest repeat (11-base pair stem, three-base loop) was found in pBR322 DNA and this potential cruciform (Fig. 5) is the site of specific cleavage.

It may be necessary for the inverted repeat sequence to be within an A+T-rich segment of much larger size to form a relatively stable cruciform. The single cleavage site in pVH51 is located in a unique, easily denaturing and unusually A+T-rich region[18]. Likewise, the specific cleavage site in pBR322 DNA is located in one of the two most A+T-rich regions in that DNA. In contrast, SV40 DNA was not specifically cleaved. The experimental and theoretical denaturation maps of this DNA reveal that the region surrounding the largest inverted repeat is not particularly A+T-rich, whereas the molecule contains three other very distinct easily denaturing regions[19].

Supercoiling is known to generate locally single-stranded regions in double-stranded DNA molecules. These regions are easily denaturing stretches of DNA, due to their high A+T content[2]. Our data suggest that cruciforms which exist in A+T-rich regions are the most stable. It is possible that the non-paired nucleotides in the A+T-rich region, which are in dynamic equilibrium with the duplex form, might facilitate the formation of a cruciform. Alternatively, if an inverted repeat is within a larger G+C-rich region, the duplex form of the DNA may be the most stable. According to this model, specific nucleases would cleave supercoiled DNA at specific sites if (at least transiently) stable cruciform structures could be formed within the most easily denaturing region. Inverted repeats present outside the most A+T-rich region would not readily form cruciforms and, therefore, would not be the first sites to be cleaved. On the other hand, if the most A+T-rich region did not contain inverted repeats, cleavage might still occur at random sites inside this region due to its single-stranded nature.

Accordingly, this model would predict that linearization of SV40 DNA by S_1 and T_7 nuclease would probably occur not at the inverted repeat but, instead, at the most easily denaturing regions. Previous studies have shown that S_1 nuclease cleaves supercoiled SV40 DNA at a preferred region, which is located at 0.45–0.55 map units and coincides with an easily denaturing region[13]. We observed that linearization of SV40 DNA by S_1 and T_7 nucleases was not reproducibly found at a specific site in the relatively low-temperature conditions which lead to position-specific cleavage of pVH51 DNA. In contrast, at high temperatures (72 °C), specific linearization was observed at position 4,540 ± 50 base pairs of the published sequence[20]. This position corresponds to 0.55 map units, in the middle of one of the three most A+T-rich regions. Thus, these results are consistent with the above predictions.

In a special case, Gellert *et al.* (reviewed in ref. 2) constructed a completely inverted repeat DNA by recombinant techniques and showed that it could assume a cruciform structure.

Conclusions

We report here the first evidence supporting the existence of cruciform structures in naturally occurring DNA sequences. The supercoiled forms of two different DNAs (pVH51 and pBR322) contain specific cleavage sites for two endonucleases (S_1 and the T_7 gene 3 product) which are specific for single-stranded regions. Both sites are in the centre of unusually long, inverted repeats and coincide with the single-stranded loops of potential cruciform structures. Five other DNA molecules (G4, ΦX174, fd, M13 and SV40), which lack long inverted repeats, are not cleaved specifically. The DNAs must be in a supercoiled form to be cleaved specifically; topoisomers with a larger number of supercoiled turns are cleaved more rapidly than those which are partially or totally relaxed. R-loop structures and one nucleotide modification system were eliminated as possible explanations. These results indicate that superhelicity induces tertiary (cruciform) structures which expose specific bases in single-stranded conformations.

This work was supported by NIH and NSF grants CA 20279 and PCM77-15033, respectively. We acknowledge the gifts of SV40 DNA by J. Mertz, the replicative forms of ΦX174, G4, fd and M13 DNA by G. Horn, and pBR322 DNA by W. Hillen. The computer programs used to search for inverted repeats and A+T-rich regions were developed by G. Horn and R. Littlewood, respectively.

Received 21 July; accepted 10 November 1980.

1. Gierer, A. *Nature* **212,** 1480–1481 (1966).
2. Wells, R. D. *et al. Prog. Nucleic Acid Res. molec. Biol.* **24,** 167–267 (1980).
3. Hsieh, T. & Wang, J. C. *Biochemistry* **14,** 527–535 (1975).
4. Vologodskii, A. V., Lukashin, A. V., Anshelevich, V. V. & Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **6,** 967–982 (1979).
5. Wang, J. C. *J. molec. Biol.* **87,** 797–816 (1974).
6. Vogt, V. *Eur. J. Biochem.* **33,** 192–200 (1973).
7. Sadowski, P. D. *J. biol. Chem.* **246,** 209–216 (1971).
8. Center, M. S. & Richardson, C. C. *J. biol. Chem.* **245,** 6285–6291 (1970).
9. Center, M. S., Studier, F. W. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **65,** 242–248 (1970).
10. Hardies, S. C. *et al. J. biol. Chem.* **254,** 5527–5534 (1979).
11. Oka, A. *et al. Molec. gen. Genet.* **172,** 151–159 (1979).
12. Bartok, K. & Denhardt, D. T. *J. biol. Chem.* **251,** 530–535 (1976).
13. Beard, P., Morrow, J. F. & Berg, P. *J. Virol.* **12,** 1303–1313 (1973).
14. Lau, P. P. & Gray, H. B. *Nucleic Acids Res.* **6,** 331–357 (1979).
15. Baas, P. D., Teertstra, W. R. & Jansz, H. S. *J. molec. Biol.* **125,** 167–185 (1978).
16. Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **42,** 77–90 (1978).
17. Scheffler, I. E., Elson, E. L. & Baldwin, R. L. *J. molec. Biol.* **48,** 145–171 (1970).
18. Patient, R. K. *et al. J. biol. Chem.* **254,** 5548–5554 (1979).
19. Funnell, B. E. & Inman, R. B. *J. molec. Biol.* **131,** 331–340 (1979).
20. Reddy, V. B. *et al. Science* **200,** 494–502 (1978).
21. Dodgson, J. B. & Wells, R. D. *Biochemistry* **16,** 2374–2379 (1977).
22. Panayotatos, N. & Wells, R. D. *J. biol. Chem.* **254,** 5555–5561 (1979).
23. Hershfield, V., Boyer, H. W., Chow, L. & Helinski, D. R. *J. Bact.* **126,** 447–453 (1976).
24. Wells, R. D. *et al. Meth. Enzym.* **65,** 327–347 (1980).
25. Panayotatos, N. & Wells, R. D. *J. molec. Biol.* **135,** 91–109 (1979).
26. Panayotatos, N. & Wells, R. D. *Nature* **280,** 35–39 (1979).
27. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
28. Godson, G. N., Barrell, B. G., Staden, R. & Fiddes, J. C. *Nature* **276,** 236–247 (1978).
29. Sanger, F. *et al. J. molec. Biol.* **125,** 225–246 (1978).
30. Beck, E. *et al. Nucleic Acids Res.* **5,** 4495–4503 (1978).

LETTERS

Galactic γ and radio synchrotron radiation

C. G. T. Haslam*, S. Kearsey†, J. L. Osborne†, S. Phillipps† & H. Stoffel*

* Max-Planck Institute for Radio Astronomy, Auf dem Hügel 69, Bonn, FRG
† Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

The galactic γ-radiation and radio synchrotron radiation contains information on the distribution of cosmic-ray particles, magnetic field and gas throughout the Galaxy. We report here a new all-sky survey of 408-MHz radio continuum which has been used to compare the distributions of radio and γ-ray emission, with the former convolved to the response function of the COS B γ-ray detector. The overall similarity of the longitude profiles suggests that the bulk of the γ-ray emission is from the interstellar medium rather than discrete sources. On a large scale in the galactic plane the two emissivities seem to be similarly distributed, although in detail the γ-ray emission is more clumpy.

γ rays with energies in the region of interest (~100 MeV) are produced by bremsstrahlung of cosmic-ray electrons of a few 100 MeV on the interstellar gas and from the decay of neutral pions, produced in the interactions with the gas of cosmic-ray nuclei with energies of a few GeV. (The inverse Compton interactions of cosmic-ray electrons are expected to give a much smaller contribution than these two processes except, perhaps, at high galactic latitudes.) If the cosmic-ray proton to electron ratio is constant, the interstellar γ-ray emissivity is then proportional to the product of cosmic-ray intensity and total gas density. There is, of course, independent information on the distribution of interstellar gas in the Galaxy so that, in principle, the distribution of cosmic rays can be obtained (the work of the past decade on this problem is reviewed in ref. 1). In practice, there are large uncertainties because the important molecular hydrogen component of the gas has to be inferred from the distribution of CO, which is itself only known in certain regions. Also, mainly because of the rather poor angular resolution of the γ-ray detectors, there is uncertainty concerning the contribution of unresolved discrete sources to the apparently diffuse γ-ray emission along the galactic plane. Additional information may therefore be usefully sought from a comparison of the γ-ray distribution with that of radio synchrotron radiation. In the frequency range from 100 MHz to 1 GHz, where there is the best combination of good angular resolution and low contamination by thermal continuum radiation, the synchrotron radiation is produced by cosmic ray electrons of a few GeV. The synchrotron emissivity is proportional to the product of the cosmic-ray electron density and approximately the square of the magnetic field intensity (B^2). The large scale distribution of B^2 may in turn possibly be related to the distribution of interstellar gas. If, for example, the γ-ray and synchrotron emissivities were in a constant ratio then B^2 would be proportional to gas density irrespective of the variation of cosmic-ray density.

Comparisons between the γ rays and synchrotron radiation have been made by several authors (see refs 2–4). They have mainly considered γ-ray data, having rather low statistical accuracy, from the SAS 2 detector (8,000 γ rays) and the synchrotron radiation data were in a form which did not make the necessary convolution easy. We now use the new all-sky survey at 408 MHz of Haslam *et al.*[5,6]. This is from observations at Effelsberg, Jodrell Bank and Parkes. The data, smoothed to a

0028-0836/81/060470—04$01.00

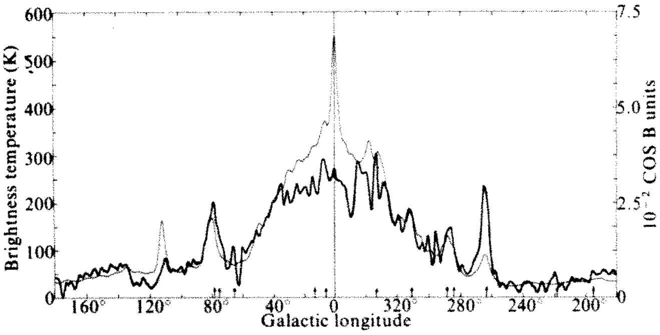


Fig. 1 Galactic plane profiles. Thick line and right-hand scale: distribution of 70 MeV to 5 GeV γ rays obtained by the COS B detector. The arrows show the longitudes of features corresponding to γ-ray 'sources'. Thin line and left-hand scale: distribution of 408 MHz continuum radiation convolved to the HPBW of the COS B detector.

common resolution of 51 arc min, are in the form of brightness temperatures at 20 arc min intervals of galactic latitude and longitude. It has been convolved and averaged appropriately for comparison with the COS B γ-ray distributions of Mayer-Hasselwander *et al.*[7] based on 64,000 photons. We present here the comparisons with latitude and longitude distributions of γ rays with energy 70 MeV $<E<5$ GeV. Comparison with the longitude distributions of γ-rays after division into three energy bands in this range leads to similar conclusions although the effects of statistical fluctuations are naturally greater.

Figure 1 shows the longitude distributions. The γ-ray intensities are given in units of 'on-axis' counts s^{-1} sr^{-1} for the COS B detector (henceforth referred to as COS B units). The thick line shows the cut along the galactic plane ($b = 0°$) of the contour map derived using the smoothing procedure detailed in ref. 7. This removes features smaller than the half-power beam width (HPBW) of the detector but not all the remaining features are necessarily statistically significant. For comparison, the 408 MHz data must be convolved to the beam profile appropriate to the γ-ray detector. Such a profile was constructed taking into account the energy dependence of the HPBW (from 7° at 70 MeV to 2° at 5 GeV) and the relative numbers of γ rays of a given energy in the COS B data: it has a HPBW of 3.5° and a longer tail than a gaussian distribution. The result of the convolution is given as the thin line curve in Fig. 1. We have chosen the cut at $b = 0°$ for the comparison rather than the longitudinal variation of the average γ-ray intensity for $|b| < 5°$, which is also given in ref. 7, because the former is affected less by the differing latitude distributions of the γ and synchrotron radiation. An indication of the statistical significance of the features in the γ-ray cut at $b = 0°$ can be obtained from the quoted statistical errors on these average intensities. We deduce that, over one beam area, statistical fluctuations of about 0.15×10^{-2} and 0.2×10^{-2} COS B units are to be expected for the outer and inner halves of the galactic plane respectively. Thus, while many of the features in the outer half of the plane can be discounted most of those in the inner half are real.

The latitude distributions are shown in Fig. 2. The points are the intensities at 1° intervals in latitude averaged over the longitude intervals indicated. The lines are the 408 MHz brightness temperatures convolved and averaged appropriately. As the object is to compare the galactic radio and γ-ray emission the isotropic extragalactic contributions have to be removed. From the 408-MHz data a 6 K isotropic component has been subtracted. The COS B intensities are those above a background. This is the sum of an instrumental background, an extragalactic component, and the estimated galactic radiation at $b = 90°$. Strictly this last component should be added back in. From the SAS 2 experiment[8] for $E > 70$ MeV this is $\sim 2.2 \times 10^{-5}$ cm^{-2} s^{-1} which seems to correspond to about 0.6×10^{-3} COS B units. For the longitudinal profile this is negligible. For the latitude profiles the addition would be just large enough to make all of the plotted points positive. Although the 408 MHz continuum radiation is predominantly non-thermal there is some thermal contribution. In the plane of the Galaxy ~20% of the extended emission is thermal. The scale height of the thermal emissivity is smaller than that of the non-thermal radiation by a factor of between 5 and 10, so that after convolution the contribution at $b = 0°$ is no more than a few per cent. The broad peak in emission underlying the sources in the Cyg X region ($l \approx 80°$) is predominantly thermal, however.

Let us now consider the latitude and longitude profiles. In Figs 1 and 2 the radio γ-ray intensities are plotted with the same arbitrary normalization ($100\,K \equiv 1.25 \times 10^{-2}$ COS B units). This makes the longitudinal profiles generally coincide (to within ~35° of the galactic centre). The most obvious difference between the profiles is that there is no strong γ-ray source at the galactic centre corresponding to the radio source Sag A. This gives a contribution of 150 K to the convolved radio profile at $l = 0°$ but has no effect on the profile for $4° < l < 356°$. Two other features must be discounted. The peak in the radio at $l = 112°$ is due to the supernova remnant Cas A ($l = 111.7°$ $b = -2.1°$); this is not a detectable source of γ rays. The peak in the γ rays at $l = 263.5°$ is essentially all pulsed emission from the Vela pulsar while the corresponding peak in the radio is unpulsed radiation from the Vela supernova remnant.

On the south side of the Galaxy the steps at 285°, 310° and 330°, usually regarded as evidence of spiral structure, are reproduced on both distributions while on the north side there are no steps in either distribution between 40° and 60° where one might expect the spiral arms to reappear. Both the distributions have a broad minimum at $l = 240°$. The rather close resemblance of the two profiles (the discrepancy within 35° of the galactic centre is discussed below) supports the proposition that the γ-ray emission, like the synchrotron radiation, is predominantly produced in the interstellar medium rather than being due to discrete sources. It is possible, of course, that such a distribution of emission could be produced by discrete sources but their number density would have to follow rather closely the variation of the product of cosmic-ray density and the square of the galactic magnetic field strength. There does, however, seem to be more structure in the γ radiation than in the radio smoothed to the same HPBW even after statistical fluctuations are taken into account. To investigate the clumpiness of the γ-ray emissivity further more detailed data than that currently published are required. Not only the maxima but also the minima in the γ-ray distribution are of interest in this context. It is possible to account for the minimum at $l = 62°$ in terms of interstellar production. This dip is almost certainly real although its sharpness may be the result of fluctuation. In this direction there is the longest line of sight to the nearest spiral arm in the whole galalctic plane. When the Perseus arm is eventually reached at a distance of 10 kpc the scale height of the gas subtends an angle much less than the beam size and its contribution to the γ-ray intensity is much reduced. The effect will be less marked for the synchrotron radiation which has a considerably larger scale height.

Catalogues of γ-ray sources have been produced from the COS B data. A source is defined as a significant excess of counts compatible with the instrument point spread function so that a feature extending to ~1° is not distinguished from a point source. The second COS B catalogue[10] contains 22 sources at low galactic latitudes. The longitudes of the 12 which contribute to features on the longitudinal profile are indicated by arrows on Fig. 1. Apart from the Vela pulsar there is no certain identification of any of these with a radio source but five features are coincident in position with radio sources strong enough to give features on the convolved radio profile. 2CG075+00 and 2CG078+00 are in the Cyg−X region. 2CG284−00 and 2CG288−00 are in the Carina spiral arm. In both regions there are various radio sources with which they might be identified including supernova remnants–OB associations (SNOBs)[11]. The fifth source 2CG006−00 is perhaps the prime candidate for a

SNOB, being positionally coincident with the bright supernova remnant W28 and the HII region RCW146 both of which appear to be at a distance of 1.6 kpc from the Sun. In the direction of increasing longitude from the peak due to this source the next four γ-ray peaks all have corresponding features in the radio profile. The completely different behaviour of the profiles on the other side of the galactic centre suggests, however, that this is a chance effect.

Comparison of the latitude profiles in Fig. 2 shows that the synchrotron emission is much broader in extent than the γ radiation. This is true for all longitude ranges except $145° < l < 175°$. In the inner part of the Galaxy, as would be expected for production in the interstellar gas, the true latitude distribution of the γ radiation is masked by the HPBW of the detector. Taking the profiles in turn one sees that in Fig. 2*b* the North Polar Spur enhances the synchrotron emission at positive latitudes but not the γ radiation. In Fig. 2*c* the blackened area shows the contribution from the bright radiogalaxy Cyg A, which is not a γ-ray source. The contribution of Cas is shown similarly in Fig. 2*d*. The γ-ray peak in Fig. 2*f* is due to the bright source 2CG195+04. There is no significant radio emission associated with it. In Fig. 2*k* the peak in the γ-ray profile is due to the Crab pulsar. The γ rays are pulsed while the radio emission (shaded) is from the supernova remnant.

The latitude profiles show that the γ-ray emissivity is not proportional to the synchrotron emissivity in three dimensions. There could still be, however, an overall proportionality between the two emissivities in the galactic plane. The deficit in the γ-ray profile within 35° of the galactic centre does not preclude this as there are two geometrical effects that must be considered. The first is a manifestation of the differing scale heights of the γ and synchrotron radiation. The reduction in the recorded intensity due to the angular width of the emissivity being much less than the beam size of the detector is larger for the narrower γ-ray emission and increases with distance. The second is the effect of a regular component of the galactic magnetic field. In the neighbourhood of the Sun a regular component of the magnetic field is observed lying in the direction of galactic rotation and of the same order of magnitude as the fluctuating component. If this azimuthal field extends throughout the Galaxy there is an enhancement in the synchrotron radiation due to its being viewed perpendicularly as the line of sight approaches $l = 0$. The distribution of synchrotron emissivity in the galactic plane derived from the full resolution synchrotron profile by an unfolding technique in which the ratio of regular to irregular field is treated as a free parameter is to be reported elsewhere[12]. A good fit is obtained for a ratio of unity and for an emissivity that peaks at a distance of 3.6 kpc from the

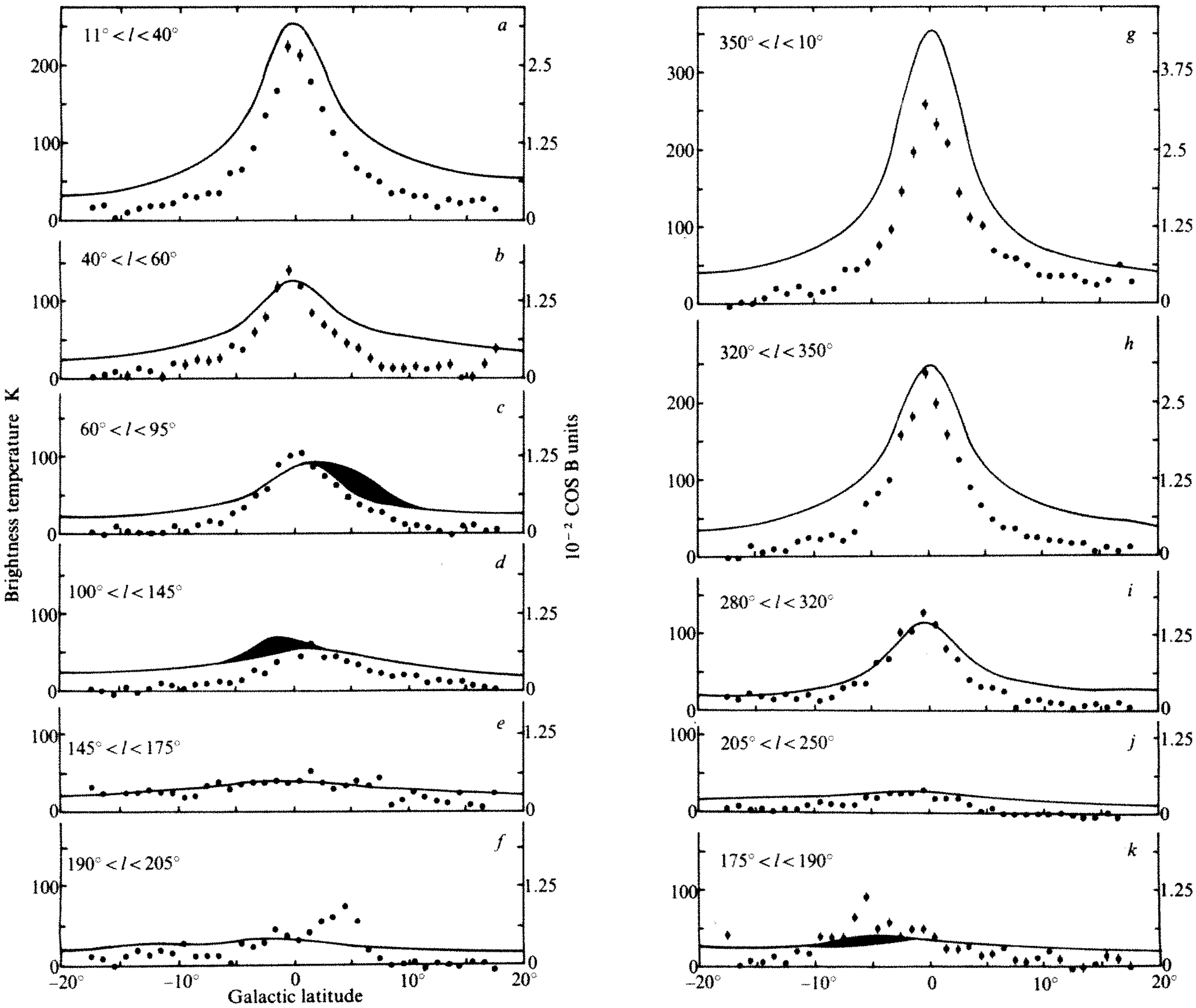


Fig. 2 Galactic latitude distributions. ●, γ-ray intensities averaged over the indicated longitude ranges. Statistical errors are shown when larger than the points. Lines, 408 MHz intensities correspondingly convolved and averaged.

galactic centre and is low within this distance. If we predict the γ-ray profile assuming that the emissivity in the plane mirrors this synchrotron emissivity but that its scale height is that of the gas, a reasonable fit to the observations is obtained. If the two emissivities are in constant ratio the implication is that B^2 is proportional to gas density on some suitably large scale in the plane of the Galaxy. Developing instabilities would extend the magnetic field out of the plane and destroy the proportionality there.

Received 2 September; accepted 5 December 1980.

1. Cesarsky, C. J. *Ann. N.Y. Acad. Sci.* **336,** 223 (1979).
2. Paul, J., Cassé, M. & Cesarsky, C. J., *Astrophys. J.* **207,** 62 (1976).
3. Jäkel, C. E., Nissen, P., Schlickeiser, R., Thielheim, K. O. *Proc. 14th int. Cosmic Ray Conf.* **12,** 4177 (1975).
4. Higdon, J. C. *Astrophys. J.* **232,** 113 (1979).
5. Haslam, C. G. T. *et al. Astr. Astrophys.* (in the press).
6. Haslam, C. G. T., Salter, C. J., Stoffel, H. & Wilson, W. E. *Astr. Astrophys. Suppl.* (in the press).
7. Mayer-Hasselwander, H. A. *et al. Ann. N.Y. Acad. Sci.* **336,** 211 (1980).
8. Fichtel, C. E., Kniffen, D. A., Hartman, R. C. & Thompson, D. J. *Proc. 12th ESLAB Symp.* (European Space Agency, Frascati, 1977).
9. Hirabayashi, H. *Publ. astr. Soc. Jap.* **26,** 263 (1974).
10. Hermsen, W. thesis, Leiden Univ. (1980).
11. Montmerle, T. *Astrophys. J.* **231,** 95 (1979).
12. Phillipps, S., Kearsey, S., Osborne, J. L., Haslam, C. G. T. & Stoffel, H. *Astr. Astrophys.* (in the press).

Laser-radar measurements in southern England of aerosols from Mount St Helens

L. Thomas, C. P. Chaloner & S. K. Bhattacharyya

SRC Rutherford and Appleton Laboratories, Ditton Park,
Slough, Berks SL3 9JX, UK

Material from the eruption of Mount St Helens (46°N, 122°W) has been detected at lower stratospheric heights using a laser-radar system at Winkfield (51°N, 1°W). The observation of a distinct layer of ~1 km thickness near 15.5 km altitude for several days in early June is consistent with the west-east transport of material from above the eruption site expected at this height and also indicates a significant northward spread of the material. In late July less pronounced layers were observed at 17-18 km and 22 km at which level the direction of motion is less certain.

The laser incorporated in the radar system operated near a wavelength of 0.6 μm and produced 1 μs pulses at a repetition frequency of $1\,s^{-1}$. The receiver included a 0.8-m diameter newtonian telescope and operated in a photon-counting mode, the return signals being recorded in channels corresponding to 1 km height intervals. The inclusion of a neutral density filter in the receiver and the restriction of the laser energy to a very low value (< 20 mJ) ensured that the photomultiplier was not overloaded by returns from low altitudes.

The results obtained near the end of May and during June are presented in Fig. 1. These show the variations with height of the range-corrected backscattered signal and standard error for ~1 h of observations on each of four nights. In each case the corresponding height variation expected for the Rayleigh-scattered signal is shown, as derived from molecular densities obtained from radiosonde data for Crawley (51°N, 0°W) provided by the Meteorological Office, Bracknell, and fitted to the observations at the lower heights. The tropopause heights indicated were from the same data. The deviations from the Rayleigh scattering results on 27 May are consistent with other measurements at Winkfield before the eruption[1]. Note the strong scattering from heights of ~15.5 km during the nights of 6 and 7 June; this feature was observed consistently between 4 and 9 June, adverse weather conditions interfering with measurements on either side of these nights. In addition, small enhancements are shown over the height range 16–24 km on the 7 June. Balloon-borne observations[2] of scattered sunlight in southern France on 5 June also showed the presence of an enhanced aerosol layer at a height near 15.5 km. The presence in the eastern US on 22 May of layers of material between 12 and 16 km from the Mount St Helens eruption was recorded with the NASA Langley airborne laser radar system[3]. Furthermore, the absence of anomalous scattering at Winkfield on 27 May was consistent with the location of the 14–16 km material at ~30°W, as observed[3] with the satellite-borne Stratospheric Aerosol and Gas Experiment (SAGE). As indicated in Fig. 1, no major scattering was observed at Winkfield during the second half of June.

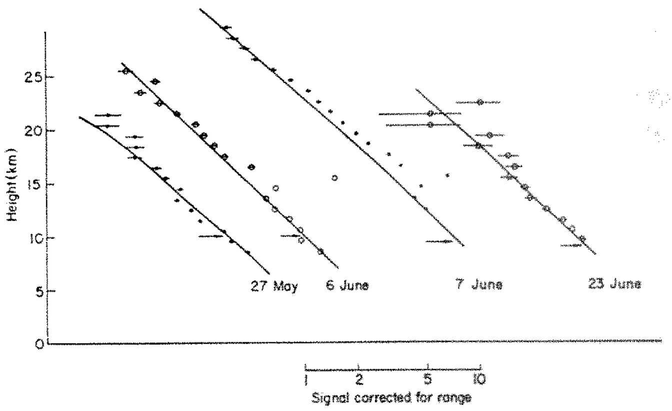


Fig. 1 Variations with height of the range-corrected signals from soundings at vertical incidence during four nights. The smooth curves represent the corresponding variations expected for Rayleigh scattering as computed for molecular densities derived from radiosonde measurements which also provided the tropopause heights denoted by arrows.

The approximate arrival date, relatively short duration and the location near 15.5 km of the layer observed at Winkfield are consistent with the transport of the layers observed previously in the eastern US by the west-east winds expected at these levels during May and June[4]. From an examination of the mean wind for the 100 mbar level, as derived by the Meteorological Office from radiosonde data, it is considered that a significant spread of the material towards the north must also have occurred. Unusual twilight effects following previous volcanic eruptions[5,6] showed that the dust injected can spread over very large areas within about 2 months of the events. Furthermore, balloonsonde[7] and laser-radar observations[8,9] revealed that the maximum enhancements of material extended over a considerable part of the 15–20 km altitude range, and continued for many months after the eruptions. Small enhancements in scattering have been observed from levels near 18 and 22 km during nights in the period 21–25 July, the only observable nights during that month.

We thank the Meteorological Office, Bracknell, for radiosonde data and information on winds at the 100 mbar level.

Received 26 September; accepted 17 December 1980.

1. Gibson, A. J., Thomas, L. & Bhattacharyya, S. K. *J. atmos. terr. Phys.* **39,** 657–660 (1977).
2. Ackerman, M., Lippens, C. & Lechevallier, M. *Nature* **287,** 614–615 (1980).
3. *NASA Upper Atmos. Prog. Bull.* No. 80.3 (1980).
4. Newell, R. E., Kidson, J. W., Vincent, D. G. & Boer, G. J. *The General Circulation of the Tropical Atmosphere* Vol. 1 (MIT Press, 1972).
5. Volz, F. E., *Tellus,* **17,** 513–515 (1965).
6. Volz, F. E. *Science* **189,** 48–50 (1975).
7. Hofmann, D. J. & Rosen, J. M. *J. geophys. Res.* **82,** 1435–1440 (1977).
8. Grams, G. & Fiocco, G. *J. geophys. Res.* **72,** 3523–3542 (1967).
9. McCormick, M. P., Swissler, T. J., Chu, W. P. & Fuller, W. H. *J. atmos. Sci.* **35,** 1298–1303 (1976).

Similarities in amorphous and crystalline transition metal–metalloid alloy structures

P. H. Gaskell

University of Cambridge, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

In many amorphous and glassy systems the local structure seems to be essentially the same in amorphous and crystalline phases. It thus becomes possible to define a 'local structural unit', such as the SiO_4 group in silicates. Determination of the structure then reduces to the problem of describing how the local units are connected and packed in the medium-range structure. There is mounting evidence that alloys of transition metals with metalloids are more adequately described by arrangements of stereochemically-defined structural units than by alternative descriptions in terms of random packing of atoms. There are three phases, Fe_3C, Ti_3P and Fe_3P, which may serve as models of local structural units. These lattices differ to a small degree in the symmetry of the metalloid nearest-neighbour shell but there are major differences in the arrangement of local structural units. Although the experimental evidence is limited, it seems that the distinctive features of the local and medium-range topologies of the crystalline phases are represented in the structures of each of the corresponding glasses.

Models for transition metal–metalloid (T–M) glasses have been constructed[1,2] by randomly connecting local structural units—trigonal prisms consisting of six T atoms centred on M atoms, Fig. 1, using the edge-sharing arrangement found in the cementite Fe_3C lattice. Crystalline alloys with small metalloid radius ratios, such as Pd–Si, Ni–B and Co–B, have this structure. Calculations based on the model agree well with high resolution measurements of neutron scattering from amorphous Pd–Si alloys[3,4], giving a quality of fit at least equal to that of any other model. Moreover the model provides a good representation of the partial pair correlation functions for this alloy. Whether this type of model adequately describes other transition metal–metalloid glasses is particularly relevant for systems such as Ni–P or Fe–B in which crystalline T_3M phases are isostructural with Fe_3P or Ti_3P lattices respectively.

A relatively broad distribution of T–M bond lengths in Fe_3C structures, Fig. 2, reflects the distinction between distances from M to vertex or capping T atoms. This is particularly noticeable when M is small and T large, for example, in Pd_3B, where the trigonal prism is almost undistorted. The metalloid radius ratio is generally larger in Fe_3P and Ti_3P lattices causing extensive dilatation of the trigonal prism and distances to the nine first neighbours become almost identical (Fig. 2).

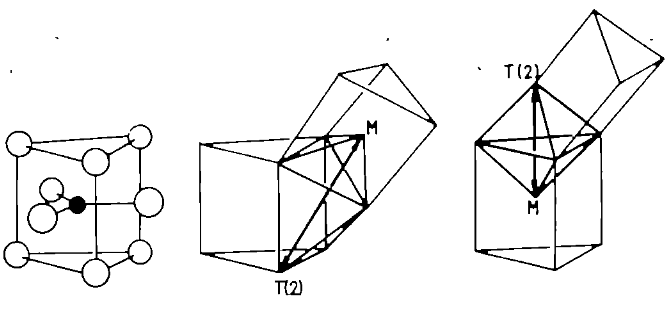


Fig. 1 Trigonal prismatic packing of transition metal atoms (large circles) around a metalloid atom observed in many crystalline alloys, together with the edge-sharing arrangement of trigonal prisms found in the Fe_3C structure (centre) and in Fe_3P (right). (Some atoms are omitted for clarity.) The two arrangements may be distinguished by distances from metalloids to second-neighbour transition metal atoms (M–T(2)). These can be visualized as the separation between two vertices of a tetrahedron sharing a face with a half octahedron (Fe_3C) and the distance between vertices of two tetrahedra (Fe_3P), the latter generally being shorter as shown.

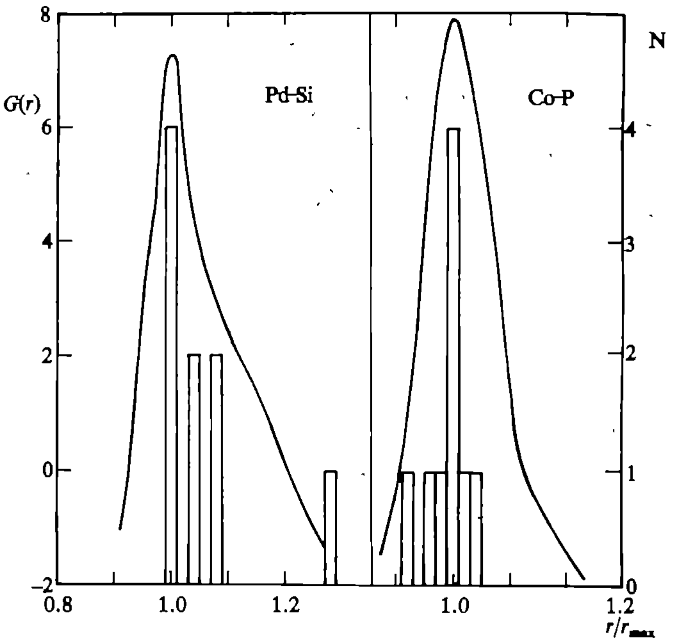


Fig. 2 Comparison of the profiles of the first-neighbour metalloid–transition metal distributions in crystalline and amorphous alloys. Bar charts represent the distribution of first neighbours (right-hand ordinate) in crystalline $Pd_3Si(Fe_3C)$ and Fe_3P lattices, as a function of reduced interatomic distance, r/r_{max}, where r_{max} is the value corresponding to the maximum of the distribution. Curves represent Si–Pd and P–Co partial pair distribution functions (left-hand ordinate) for glassy $Pd_{84}Si_{16}$ and $Co_{81}P_{19}$ alloys[5,6].

Extraction of T–M partial pair distribution functions for amorphous alloys requires an analysis of three essentially independent scattering measurements (for example, using X-ray and polarized neutrons) and this is rarely possible. Two good examples exist: $Pd_{84}Si_{16}$ (ref. 5) and $Co_{81}P_{19}$ (ref. 6). The former is potentially a 'Fe_3C glass' and the latter arguably related to Fe_3P on the basis of its radius ratio and crystallization products in Co–P–B ternary alloys. The marked differences between the profiles of the first Si–Pd and P–Co peaks is shown in Fig. 2. Furthermore, the T–M distribution for each glass is clearly similar to the bar chart for the corresponding crystalline phase. This strongly suggests that the environment of the non-metallic atoms is not random but is predominantly trigonal prismatic as in crystals. Moreover, distortions of the trigonal prisms resulting from the central metalloid atoms seem to be essentially identical in each type of amorphous and crystalline compound, suggesting a degree of stability for this type of coordination.

The disposition of trigonal prisms in the three crystalline lattices may be characterized as packing of trigonal prisms around half octahedra in Fe_3C or around tetrahedra in Ti_3P or Fe_3P lattices (Fig. 1) and it is possible to distinguish each type of packing through differences in the distances from a metalloid to second neighbour metal atoms. Second neighbour M–T(2) distances are shown in Fig. 1 and correspond to the vertex–vertex separation of a distorted tetrahedron sharing a triangular face with a half octahedron (Fe_3C) and the distance between vertices of two tetrahedra sharing a triangular face (Ti_3P and Fe_3P).

Second neighbour distances can be calculated for each pair of polyhedra as a function of the metalloid radius ratio, $p = r_M/r_T$. For the bi-tetrahedral configuration:

$$r_{M-T(2)} = (p^2+2p-\tfrac{1}{3})^{1/2}+(8/3)^{1/2}$$

and for the tetrahedral–octahedral configuration:

$$r_{M-T(2)} = [p^2+2p+5-4(p^2+2p)^{1/2}\cos\beta]^{1/2}$$

0028-0836/81/060474—03$01 00

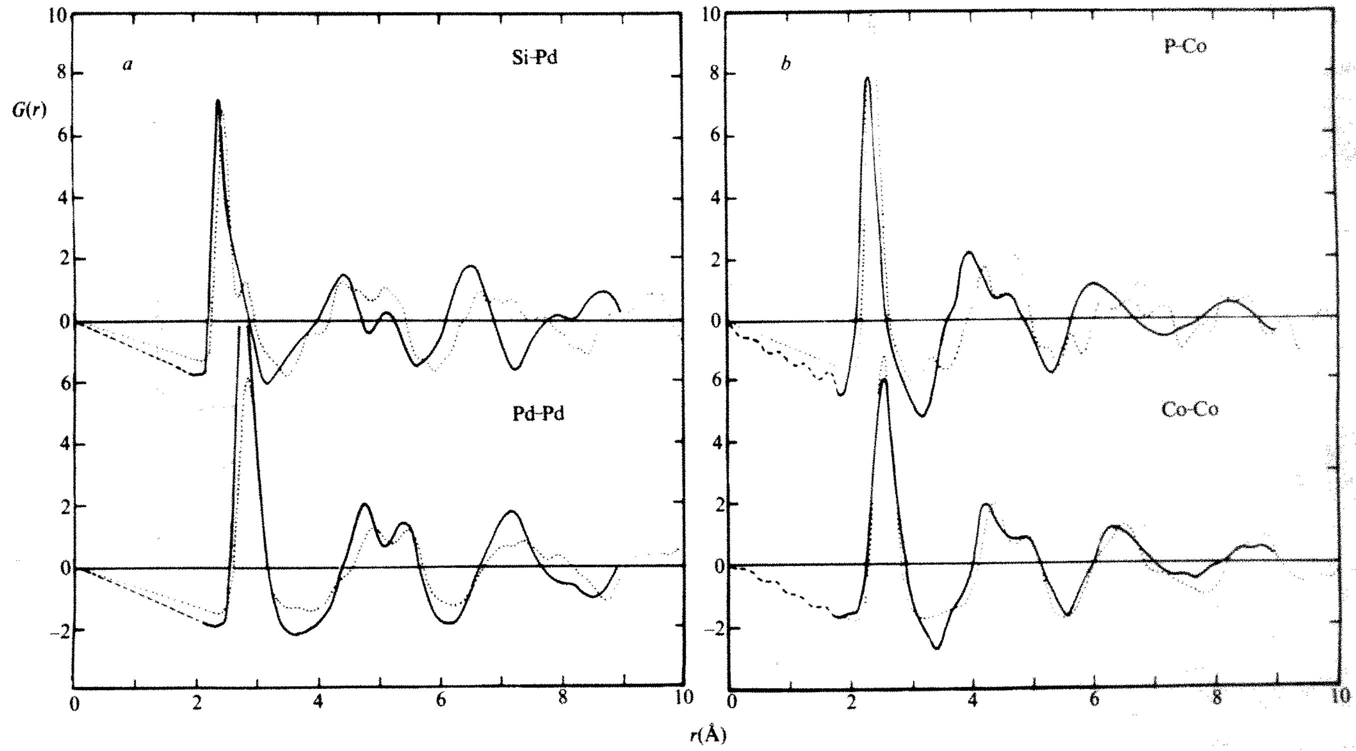


Fig. 3 Experimental (solid lines) and calculated (dotted lines) partial pair distribution functions for an amorphous Pd–Si alloy (*a*) and a Co–P alloy (*b*). Experimental data from Sadoc and Dixmier[5,6]. Si–Pd correlation functions have been derived from two experimental measurements only: an approximation which should be adequate because Si–Si correlations are very weak and the form of this distribution can be assumed with sufficient accuracy.

where $\beta = \psi_1 + \psi_2$. The angles ψ_1 and ψ_2 are dihedral angles for a half octahedron and a distorted tetrahedron,

$$\psi_1 = \cos^{-1}(3^{-1/2}), \quad \psi_2 = \cos^{-1}[(3p^2+6p)^{-1/2}]$$

The radius ratio for Co–P is 0.83 (using Sadoc and Dixmier's values[5]) suggesting a second neighbour distance, $r_{\text{M-T(2)}} =$ 3.87 Å for the bi-tetrahedral model which agrees well with the experimental value, 3.95 Å. The corresponding tetrahedral–octahedral distance is 4.14 Å. Conversely, for $Pd_{84}Si_{16}$ ($p =$ 0.74) the best fit to experiment, $r_{\text{M-T(2)}} = 4.40$ Å, is given by the tetrahedral–octahedral model, 4.35 Å compared with 4.05 Å for the bi-tetrahedral arrangment. It seems, therefore, that the medium-range structure in these two amorphous alloys is essentially different and that there may be some equivalence of the medium-range topologies of crystalline and amorphous lattices.

This calculation may be confirmed by computer-simulation. The aperiodic model constructed with medium-range topology equivalent to Fe_3C (refs 1, 2) gives good agreement with the total neutron structure factor of amorphous Pd_4Si and the total radial distribution function. As Fig. 3*a* shows, this model also gives an excellent representation of the Pd–Pd and Si–Pd partial correlation functions obtained by Sadoc and Dixmier[5]. A split peak in the computed Si–Pd distribution is introduced by setting values for the equilibrium Si–Pd (vertex) distance of 2.4 Å, and 2.8 Å for the distance to capping atoms. This is a crude model of the nearest-neighbour arrangement shown in Figs 1 and 2. Both components of the second neighbour peak Si–Pd at 4.4 and 5.1 Å are satisfactorily reproduced, although the position of the first and the intensity of the second are somewhat inaccurate.

The model has been rescaled for Co–P without making any essential alterations to the topology, that is trigonal prisms and Fe_3C-type interconnections persist. Changes in nearest neighbour geometry have been introduced in an attempt to represent the local structure in Fe_3P-type alloys, specifically, the distinction between distances to vertex and capping atom mentioned above has been withdrawn. Computed Co–Co and P–Co pair correlation functions are shown in Fig. 3, together with measurements of Sadoc and Dixmier[6]. Co–Co distances are again reproduced adequately to ~8 Å so that the total radial distribution function and structure factor, which are weighted by the Co–Co distribution, agree well with experiment[7]. The profile of the first P–Co peak and the position of the second component of the second neighbour peak at ~4.9 Å agrees with experiment, but the first component at 3.95 Å is inadequately represented by a peak at 4.17 Å in the computed data. Several attempts have been made, by altering values of the cohesive energy, changing equilibrium interatomic distances, to reduce this discrepancy and at the same time preserve a good fit to the rest of the data, but none has been successful. This difference, although small, seems to be significant and represents a closest P–Co(2) distance in an essentially tetrahedral–octahedral arrangement. Although a full model, built on packing principles exhibited by Fe_3P lattices, is required to prove the point, it seems that the bi-tetrahedral arrangement shown in Fig. 1 will be necessary to model successfully scattering data for cobalt phosphorus alloys.

Further experimental investigations of the partial correlation function of other amorphous alloys corresponding to Fe_3C and Fe_3P crystalline phases are needed and while the data, for which the errors have not been established, are sparse, current information suggests two conclusions. First, amorphous T–M glasses seem to contain a common local structural unit in which the detailed arrangement is basically determined by properties, such as the radius ratio, of the two elements. The local structure, therefore, is essentially identical to that of a compositionally-related crystalline phase. Second, medium-range structures for different combinations of elements are determined by rules, which are as yet incompletely deciphered, although the radius ratio again seems to be a dominant parameter. The result is that Fe_3C- and Fe_3P-type amorphous alloys (and crystals) possess essentially dissimilar medium-range topologies. This seems to be the first example of detectable medium-range structural differences in amorphous alloys.

Although there is little direct evidence, the composition-dependence of some magnetic properties of crystalline and amorphous Fe–(P, B) alloys[8,9] indirectly supports the idea of differences in medium-range structures. For alloys with B : P ratio of ≤ 0.5 the stable crystalline form is Fe_3P, otherwise Ti_3P

(except for compositions close to Fe_3B where Ti_3P is metastable). The composition-dependence of the Curie temperature of both crystalline and amorphous alloys shows breaks corresponding to the transition between crystalline forms. These measurements have been interpreted in terms of changes in short-range order but the local structure is almost identical in the two crystalline forms. Therefore, even if changes of magnetic properties are directly related to local structure, these are probably accompanied and perhaps driven by changes in the mode of packing of prismatic units.

Recently, evidence from NMR measurements[10] of the similarity in local structure in crystalline and amorphous alloys and of the dependence of local structure on composition has been reported. The environment of boron in glassy $Ni_{78}P_{14}B_8$ is found to be similar to that of Ni_3B which has the Fe_3C structure. In $Mo_{48}Ru_{32}B_{20}$, boron has cylindrical local symmetry which could reflect an essential similarity between boron coordination in the glass and in crystalline Mo_2B in which the atom is surrounded by eight Mo atoms at the corners of a square antiprism. However, as crystalline Fe_3P is also tetragonal, the NMR results are consistent with this type of structure also (Durand, personal communication).

Finally, local and medium-range topologies and the question of microcrystallinity are essentially distinct facets of any structural description of amorphous solids. The information presented here seems to be inconsistent with simple hard-sphere packing models for amorphous alloys; equally it is unnecessary to invoke any form of periodicity.

This work has been carried out during the tenure of an SRC Senior Fellowship. I thank the SRC and Pilkington Brothers Ltd for their support.

Received 16 October; accepted 26 November 1980.

1. Gaskell, P. H. *Nature* **276,** 484–5 (1978).
2. Gaskell, P. H. *J. Non-Cryst. Solids* **32,** 207–24 (1979).
3. Suzuki, K., Fukunaga, T., Misawa, M. & Masumoto, T. *Science Rep. Inst. Tohoku Univ.* **26A,** 1–11 (1976).
4. Fukunaga, T., Misawa, M., Fukamichi, K. & Suzuki, K. in *Rapidly Quenched Metals III* Vol. 2 (ed. Cantor, B.) 325–329 (Metals Society, London, 1978).
5. Sadoc, J. F. & Dixmier, J. in *The Structure of Non-Crystalline Materials* (ed. Gaskell, P. H.) 85–88 (Taylor & Francis, London 1977).
6. Sadoc, J. F. & Dixmier, J. *Mater. Sci. Engng.* **23,** 187–92 (1976).
7. Chi, G. C. & Cargill, III, G. S. *J. appl. Phys.* **50,** 2713–20 (1979).
8. Durand, J. & Yung, M. in *Amorphous Magnetism II* (eds Levy, R. A. & Hasegawa, R.) 275 (Plenum, New York, 1977).
9. Luborsky, F. E. in *Ferromagnetic Materials* Vol. 1 (ed. Wohlfahrt, E. P.) 451–530 (North-Holland, Amsterdam, 1980).
10. Panissod, P. *et al. Phys. Rev. Lett.* **44,** 1465–68 (1980).

Formation and behaviour of coal free radicals in pyrolysis and liquefaction conditions

L. Petrakis & D. W. Grandy

Gulf Research and Development Company, PO Drawer 2038,
Pittsburgh, Pennsylvania 15230

In the past several years there has been a renewed interest in fundamental studies of coals and coal conversion. This interest is due to the expectation that an understanding of the structure of coal and of the fundamental processes during coal conversion would aid the development of technology for the utilization of coal on cost-effective and environmentally acceptable terms. In this context, electron spin resonance (ESR) is a promising technique, for it can be used to study paramagnetic centres in coal, including reactive free radicals which are generally assumed to have a key role in conversion processes. The ESR measurements of coals reported to date have all been measurements of pre-existing free radicals or free radicals remaining after some treatment. We report here the fabrication and utilization of a uniquely designed high-pressure–high-temperature ESR cavity and the first *in situ* observation of free radicals in coals.

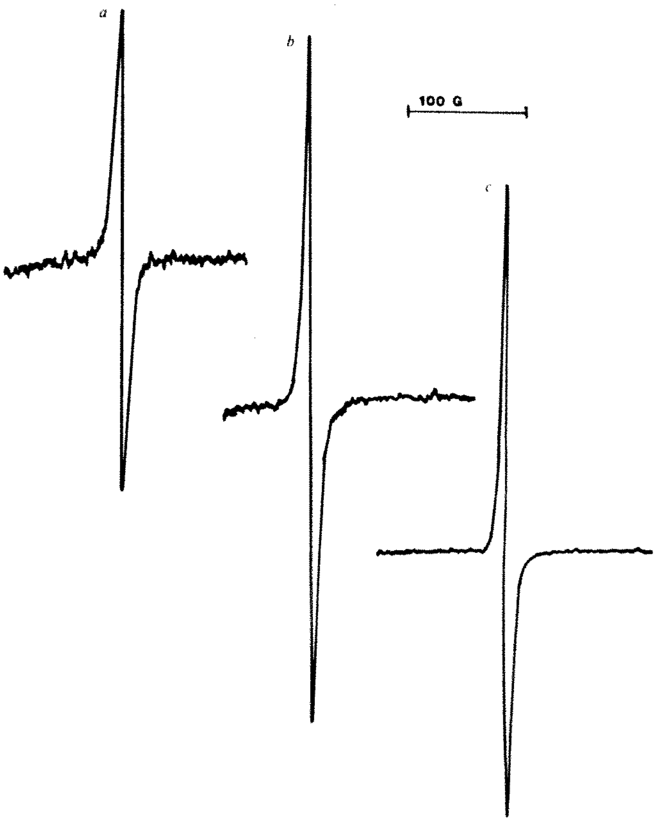


Fig. 1 ESR spectrum of Powhatan no. 5 bituminous coal under 10.335 MPa of hydrogen. Sample temperatures: *a*, 20 °C before heating; *b*, 450 °C; and *c*, 20 °C after 450 °C heating. Relative spectrometer gain factors are 8, 2.5 and 1, respectively.

This development makes it possible to perform a wide range of experiments such as coal conversion studies in process-relevant conditions, coke formation and various catalytic reactions. Findings from the *in situ* monitoring of the formation and behaviour of coal free radicals include: residual free radical concentrations following pyrolysis are similar to concentrations at reaction temperatures; manipulation of typical process variables effect a 12-fold increase in radical concentration, while the experimental error is much smaller (±40%); the determination that the most important process variable is temperature; process-derived solvents behave quite differently from single-component hydro-aromatic solvents such as tetralin.

Naturally occurring free radicals in coals were first reported in 1954. Subsequent studies have shown that the characteristics of pre-existing free radicals depend on the quality of the coals[3], that they offer the possibility of probing the structure of coal[3–5], and that additional free radicals may be generated on various thermal treatments[2,5–8]. This latter observation is particularly interesting, as it is generally assumed that stabilization of the free radicals through hydrogen abstraction from hydrogen-

Table 1 Comparison of yields in conversion of Powhatan no. 5 in ESR experiments

Solvent	Overall conversion (%)	Oil yields (%)
SRC-II*	27–36	2–6
Tetralin*	32–58	12–40

* 460 °C, low residence time.

0028-0836/81/060476—02$01.00

donating solvents may be an important aspect of coal liquefaction[9]. Alternatively, recombination of free radicals at high temperatures could result in undesirable, high-molecular-weight species that cause reactor problems. Thus, it would be useful to be able to study the formation and/or reactions of free radicals *in situ*, that is in high-temperature and high-pressure conditions. The detailed design and operation of our ESR cavity that allows such measurements are discussed elsewhere[10]. A different high-pressure ESR application has also been described recently[11].

Figure 1 shows spectra that have been obtained from a Powhatan no. 5 HVA bituminous coal. This coal is one that has been used in pilot plant studies for the Solvent Refined Coal-II (SRC-II) process. Figure 1*a* is the first-derivative spectrum of the free radicals of natural coal at room temperature and 10.335 MPa pressure of hydrogen. Figure 1*b* is the signal obtained while the same sample is held at 450°C and the pressure of hydrogen in the sample is at 10.335 MPa. Finally, Fig. 1*c* shows the same sample after termination of the treatment and re-establishment of room temperature at hydrogen pressure of 10.335 MPa. By recording reference spectra, spin concentration and *g*-values can be readily determined, as well as how they may change with various treatments and time. On thermal treatment (450 °C), the free radical concentration of Powhatan coal increases several-fold from 8.6 ± 1.0 to $59 \pm 5.0 \times 10^{18}$ spins per gramme. The observed increase is expected and it may be even greater for other types of coals and coal components[5]. On returning to room temperature, the observed residual free radical concentration is about the same ($53 \pm 5.0 \times 10^{18}$ per g) as had been observed at the high temperature.

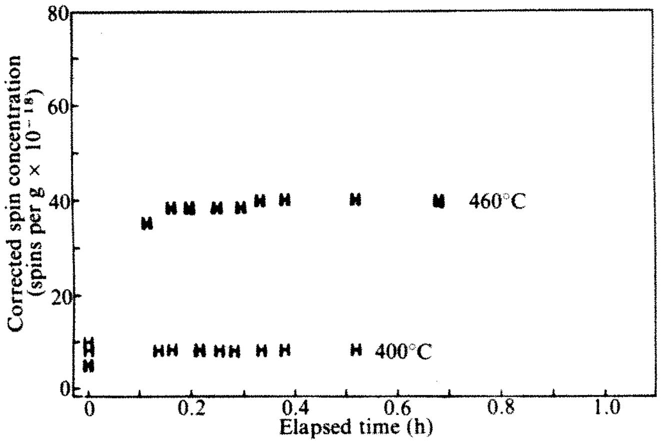


Fig. 2 Spin concentration plotted against elapsed time for Powhatan no. 5 and Tetralin. 1,200 p.s.i. H_2; HT = 3.

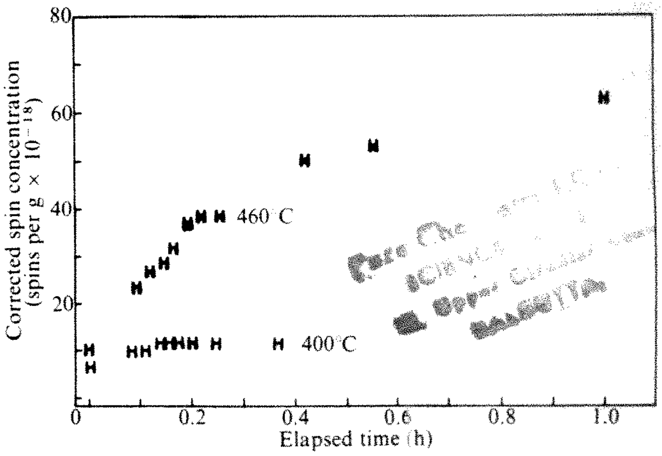


Fig. 3 Spin concentration plotted against elapsed time for Powhatan no. 5 and SRC-II HD. 1,200 p.s.i. H_2; HT = 3.

In addition to these pyrolysis experiments, more than 60 experiments, including a partially replicated 2^5 factorial design, have been carried out to assess: (1) the role of liquefaction-process variables, such as solvent, temperature, gas, pressure and heating rate, on the formation of free radicals; and (2) how the degree of conversion and quality of product are correlated with the measured free radical parameters. ESR spectra can be obtained continually starting 3 min after the initiation of the heating cycle. Figures 2 and 3 show the variation of the free radical concentration as a function of residence time. These results allow the effects of temperature (400 and 460 °C) to be assessed with two different solvents (tetralin and the heavy distillate of the SRC-II process-derived solvent from the Powhatan no. 5 coal). All other experimental conditions are identical in the four experiments shown (that is H_2 gas at 8.37 MPa and 3-min heating time to achieve reaction temperature). With tetralin as a solvent (Fig. 2), the free radical concentration remains constant at 400 °C as a function of residence time, while at 460 °C it reaches its maximum value within 6 min. With SRC-II heavy distillate as solvent (Fig. 3) there is a small increase in free radical concentration at 400 °C, while at 460 °C the free radical concentration continues to increase even after 1 h of reaction time. The free radical concentration with tetralin as the solvent is lower than with the SRC-II heavy distillate. This is not unexpected, as tetralin is considered a good hydrogen donor and, therefore, can quench efficiently the coal free radicals.

We have also attempted to obtain data on the degree of conversion and quality of the product obtained. Such data are necessary to establish whether free radical chemistry determines the reaction paths in the liquefaction of coal[9]. The preliminary data available indicate that the overall tetrahydrofuran soluble conversion is lower when SRC-II distillate is the solvent (27–36% conversion) compared with the corresponding tetralin experiments (32–58%). For the same ESR experiments, the oil yields (pentane solubles) with SRC-II heavy distillate solvent are considerably lower (2–6%) compared with the corresponding tetralin runs (12–40%) (Table 1). These different degrees of conversion as well as the differences in the time-dependent spectra (Figs 2, 3) clearly indicate that some significantly different chemistry is taking place in the SRC-II process solvent experiments relative to the tetralin experiments. This is especially true in the first half-hour of the liquefaction reaction. The results also indicate very strongly the importance of both individual and combination effects of the process variables on the formation and behaviour of free radicals. Finally, the results shown here indicate the potential of using *in situ* measurements for understanding coal conversion processes on a molecular level, and thereby controlling the quality of the liquefaction product through the manipulation of the free radical chemistry by the various process variables.

This work has been performed partly under US Department of Energy contract 01-79ET14940.

Received 10 January; accepted 24 November 1980.

1. Uebersfeld, J., Etienne, A. & Combrisson, J. *Nature* **174**, 614 (1954).
2. Ingram, D. J. E., Tapley, J. G., Jackson, R., Bond, R. L. & Murnaghan, A. R. *Nature* **174**, 797–798 (1954).
3. Retcofsky, H. L., Stark, J. M. & Friedel, R. A. *Analyt. Chem.* **40**, 1699–1704 (1968).
4. Miyagawa, I. & Alexander, C. Jr *Nature* **278**, 40–41 (1979).
5. Petrakis, L. & Grandy, D. W. *Analyt. Chem.* **50**, 303–308 (1978).
6. Petrakis, L. & Grandy, D. W. *Fuel* **59**, 227 (1980).
7. Petrakis, L. & Grandy, D. W. *Discuss. R. Soc.* (in the press).
8. Smidt, J. & VanKrevelen, D. W. *Fuel* **38**, 355–368 (1959).
9. Wiser, W. H. in *Scientific Problems of Coal Utilization* 219–236 (National Technical Information Center, Springfield, 1978).
10. Grandy, D. W. & Petrakis, L. *J. magn. Res.* **41** (in the press).
11. Livingston, R., Zeldes, H. & Conradi, M. S. *J. Am. chem. Soc.* **101**, 4312 (1979).

Determination of geomagnetic archaeomagnitudes from clay pipes

K. P. Games & M. E. Baker*

Department of Geophysics, Oliver Lodge Laboratory,
The University of Liverpool, Oxford Street,
Liverpool L69 3BX, UK

Archaeomagnitude determinations of a selection of clay pipes dateable to AD 1645±10 as well as studies of pottery samples from the same site and of the same age have been made. Values of the magnitude of the ancient magnetic field ($\mathbf{B}_{\mathrm{anc}}$), were obtained from two pottery sherds, two pipe bowls and three pipe stems. The values from the sherds and bowls agree within 2% and compare well with the average value of the magnitude of the magnetic field for the seventeenth century as determined by other archaeomagnetic studies. However, the pipe stems give values of $\mathbf{B}_{\mathrm{anc}}$ which are significantly less than those from the bowls and pottery. We have not yet been able to explain this and thus we suggest that reliable archaeomagnitude determinations can be made from the bowls of clay pipes but not from the stems. Nevertheless, this result provides a new source of material for investigating variations in the geomagnetic field strength over the past 400 yr. Clay pipes have been manufactured in England since the end of the sixteenth century. In the firing process some pipes were broken and disposed of without ever having been smoked. One such collection, discovered at Rainford, Lancashire, in 1978, consisted of a series of discrete dumps including pipes, kiln debris and a small collection of contemporary used earthenware sherds. The internal consideration of the dumps suggested a very short period of activity and archaeologists (P. Davey, personal communication) ascribe all the material to the period 1645±10 yr. With such well-dated material, we set out to check whether or not reliable archaeomagnitudes could be obtained from the pipes.

The samples were all analysed using Shaw's[1] ARM technique. The maximum size of the samples used was determined by the sample holder which was designed to hold 1-inch diameter cylindrical cores. Figure 1 shows the graphs obtained for one of the pipe bowls. The points labelled R are rejected because they do not lie on a straight line of gradient = 1.00 on the plot of ARM 1 against ARM 2. The value of $\mathbf{B}_{\mathrm{anc}}$ is determined from the slope of the normal remanent magnetism (NRM) against thermoremanent magnetism (TRM) graph using the equation

$$\frac{\mathrm{NRM}}{\mathrm{TRM}}=\frac{\mathbf{B}_{\mathrm{anc}}}{\mathbf{B}_{\mathrm{lab}}}$$

where $\mathbf{B}_{\mathrm{lab}}$ is the laboratory magnetic field in which the TRM was acquired, in this case 50μT.

All the samples were given their TRMs over a period of 7 h to reproduce as closely as possible the original slow cooling rate. This minimizes the error introduced by different cooling rates for the NRM and TRM as described by Fox and Aitken[2]. All the TRMs were installed using a field in the direction of the NRM to produce a TRM in the same direction as the original NRM within about 5°. This was to avoid the possible source of error caused by anisotropy[3]. This was very important as some of the samples were very anisotropic (see Table 1).

The results obtained for the pottery, bowl and stem samples are given in Table 1. To ascertain how close these results were to the actual value of the magnetic field at that time we averaged all of the known archaeomagnitude determinations for the seventeenth century as published by the Academy of Sciences (USSR)[4]. Seventy-two results for this period have an average value for $\mathbf{B}_{\mathrm{anc}}/\mathbf{B}_0$ of 1.12 ± 0.13 where $\mathbf{B}_0$ represents the present day field value at the place where the sample came from, and where the error is the r.m.s. deviation from the mean. For Rainford, therefore, we would expect the value of the ancient magnetic field strength to be 53.8 ± 6.2 μT. Table 1 shows that the values of $\mathbf{B}_{\mathrm{anc}}$ from the pottery and bowl samples agree very well with each other and with the expected value. Averaging both sets of results we get 51.0 ± 2.0 μT for the bowls and 52.0 ± 1.0 μT for the pottery. However, the three pipe stems studied give widely ranging values which are all less than the expected value. Two of these stems, R7810 and R7813, give values which are of the order of 30% and 15% too low compared with the pottery and bowls, while the third stem, R7814, gives a value which is about 5% too low.

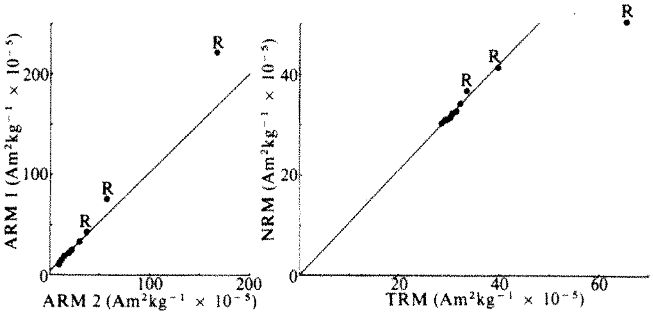


Fig. 1 Results for bowl sample R 7822. The graphs show all of the data for the NRM and TRM and ARM 1 versus ARM 2 plots. The points labelled R are rejected because they do not lie on the straight line of gradient = 1.00 drawn on the plot of ARM 1 versus ARM 2.

We have not yet succeeded in finding the reason for these low values from the stems, but we determined the degree of TRM anisotropy present in all of our samples, and these values are given in Table 1. Although two of the stems show the greatest degree of anisotropy, this in itself should not affect the value of $\mathbf{B}_{\mathrm{anc}}$ as the TRM was always installed along the same direction as the NRM. Moreover, the third stem actually has quite a small degree of anisotropy.

These pipes were possibly placed in a radial pattern around a central pillar inside a cylindrical oven, with the bowls on the outside and the stems towards the centre. The lengths of the

Table 1 Magnetic field measurements of pottery bowl and stem samples

Sample	Type	NRM intensity ($\mathrm{A\,m^2\,kg^{-1}} \times 10^{-5}$)	TRM anisotrophy (%)	$\mathbf{B}_{\mathrm{anc}}$(μT)	Average value of $\mathbf{B}_{\mathrm{anc}}$(μT)	Angle between NRM and long axis of stem
R 7811	Bowl	110	6	49±2	51±2	—
R 7822	Bowl	50	14	53±3		—
R 7820	Pottery	14	6	51±3	52±1	—
R 7823	Pottery	91	10	53±3		—
R 7810	Stem	65	34	36±5	42±5	85°
R 7813	Stem	36	7	42±1		25°
R 7814	Stem	134	26	48±3		10°

---

* Present address: Geophysical Service International, Manton Lane, Bedford, UK.

0028-0836/81/060478—02$01.00

pipes may have been longer than the radius of the ovens, and so there would be a certain amount of stacking and overlapping of pipe stems in the central region. We calculated the effect of the distortion of the field produced by the overlapping stems, and concluded that it was too small to account for the low values of the field which we have obtained.

Another way in which the pipes could have been arranged was with the bowls at the outside of the oven, with the stems pointing radially inwards but also raised vertically so that the ends of the pipe stems met at some common point. As the inclination of the magnetic field is quite steep in England, this configuration would mean that the direction of the NRM in the pipe stems would in some cases be nearly parallel to the length of the stem and in other cases it would be transverse to the stem, whereas in the previously described configuration there is no obvious way that the NRM can be acquired along the length of the stem. When we analysed the NRM directions in the stems, we found that the direction of the NRM varied from angles of 10° to 85° with respect to the axis along the length of the stem (see Table 1). Furthermore, there was a correlation between this angle and the value of $\mathbf{B}_{\text{anc}}$, but we cannot find any explanation for this.

We conclude, therefore, that the bowls of clay pipes may be used to determine archaeomagnitudes. As one of the major problems in archaeomagnetism is the availability of well-dated suitable material, any new source material such as these clay pipes is of great interest. Also, once the changes in the strength of the geomagnetic field over the past 400 yr have been satisfactorily determined, we may be able to date pipes of unknown age by this technique.

We thank P. Davey for samples and for his interest and help. K.P.G. is supported by NERC grant no. GR3/2396.

Received 24 September; accepted 10 November 1980.

1. Shaw, J. *Geophys. J. R. astr. Soc.* **39**, 133-141 (1974).
2. Fox, J. M. W. & Aitken, M. J. *Nature* **283**, 462-463 (1980).
3. Rogers, J., Fox, J. M. W. & Aitken, M. J. *Nature* **277**, 644-646 (1979).
4. *Archaeomagnetic Determinations of the Geomagnetic Field Elements* (Soviet Geophysical Committee of the Academy of Sciences of the USSR, 1977).

Thermoluminescence dating of late Devensian loesses in southern England

A. G. Wintle

The Godwin Laboratory, Sub-department of Quaternary Research,
Free School Lane, Cambridge CB2 3RS, UK

Scattered across southern England are many isolated deposits of loess-like material[1]. A few, such as that at Pegwell Bay in Kent, are highly calcareous and unweathered but most have been reworked by fluvial or colluvial processes. There is good stratigraphical evidence for a few pre-Devensian loesses, also in Kent, but dating of more recent loess has so far been based on indirect evidence. Much work has been done on the Pegwell Bay loess as it is the most extensive, truly aeolian loessic deposit in Britain. Kerney[2] compared the late Devensian deposits in the Isle of Thanet and at Pegwell Bay with similar deposits in Holland and Belgium where radiocarbon dates have been obtained for interstadial deposits. Correlation of the East Kent deposits with these in Northern Europe indicates that the loesses in Kent were formed between 30,000 and 14,000 yr ago. I report here dates for six of the more recent deposits in southern Britain from the Scilly Isles to Kent. The dates have been obtained on the loess itself, using a recently developed thermoluminescence (TL) dating technique[3,4], and confirm the ages as being late Devensian.

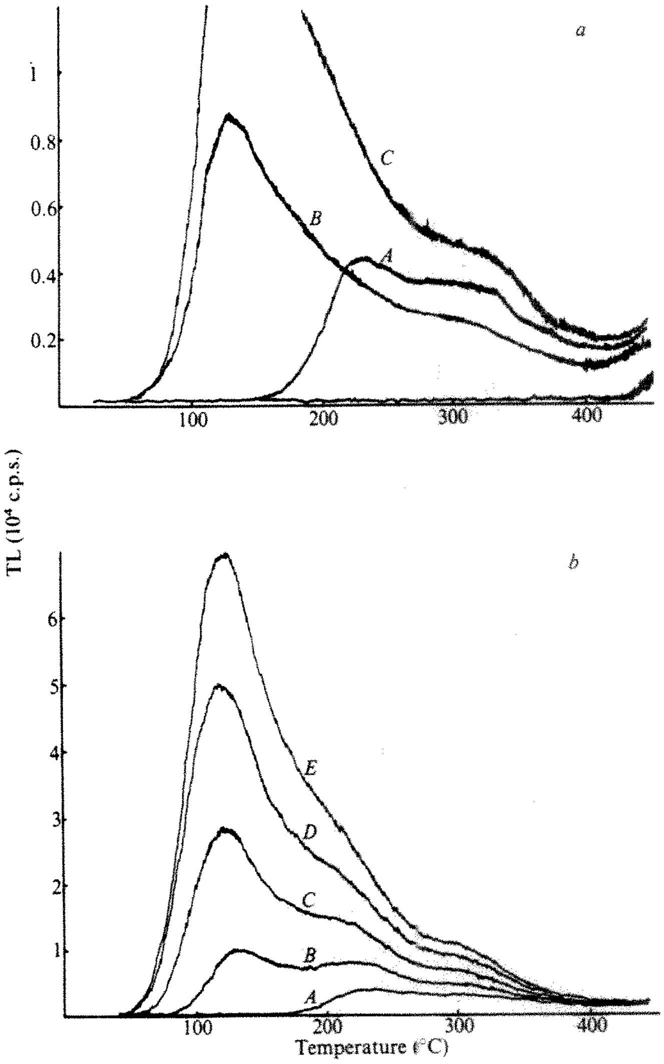


Fig. 1 TL glow curves for late Devensian loess. *a*: *A*, natural TL; *B* and *C*, TL due to 6.7 and 13.4 krad after heating. *b*: *A*, natural TL; *B–E* natural + 8.9, 13.4 and 17.9 krad respectively.

The TL dating of any sedimentary deposit, whether marine or terrestrial, is based on the assumption that exposure to sunlight during the weathering and transport of the detrital grains is sufficient to remove most of their previously acquired TL signal. The TL signal in a mineral is due to the untrapping of electrical charges in the crystal when it is heated, the trapped charges having been produced by ionization due to the decay of natural radioactive elements in the sediment. The most common TL-sensitive minerals are quartz and feldspar and these are the dominant minerals present in loess. In southern England the peak grain size of the deposits decreases from ~45 μm in Kent to 25 μm in the south-west[1] suggesting that the loesses, if contemporary, have been derived from a source in the North Sea. In this study polymineral grains in the range 4–11 μm were used, 0.5-mg samples being deposited onto 1-cm diameter aluminium disks for TL measurements. The disk is heated in an oxygen-free atmosphere from room temperature to 450 °C at 2 °C s^{-1} and the resulting TL output recorded as a function of temperature known as a glow curve. The glow curves for all six samples are very similar to each other and to those obtained for detrital grains from ocean sediments[3].

Only the Pegwell Bay sample 1a could be described as pure aeolian loess. The St Agnes sample 1d looked pure but did not

0028-0836/81/060479—02$01.00

Table 1 Thermoluminescence data and radioactivity analyses

Sample QTL no.	ED+I (krad)	% K_2O	Total α-count rate ($ks^{-1}\ cm^{-2}$)	Th* (p.p.m.)	U* (p.p.m.)	a (± 0.05)	Δ_{sat}	Dose rate (rad yr^{-1})	TL age (kyr)
1a†	$5.0+0.0=5.0$ ± 0.4	1.77 ± 0.02	0.669 ± 0.013	9.7 ± 0.8	2.6 ± 0.2	0.084	0.50	0.337	14.8
1b	$5.8+0.6=6.4$ ± 0.5	1.90 ± 0.01	0.992 ± 0.012	16.0 ± 1.5	3.0 ± 0.4	0.092	0.49	0.441	14.5
1c	$5.9+0.3=6.2$ ± 0.4	1.81 ± 0.01	0.665 ± 0.013	8.8 ± 1.2	3.0 ± 0.3	0.075	0.48	0.330	18.8
1d	$6.4+0.4=6.8$ ± 0.5	2.23 ± 0.04	0.653 ± 0.006	9.7 ± 0.9	2.6 ± 0.3	0.085	0.26	0.365	18.6
1e	$6.1+0.0=6.1$ ± 0.4	1.67 ± 0.00	0.835 ± 0.013	9.8 ± 0.8	4.1 ± 0.2	0.089	0.48	0.383	15.9
1f	$6.2+0.0=6.2$ ± 0.4	1.71 ± 0.02	0.670 ± 0.011	7.6 ± 0.9	3.4 ± 0.3	0.085	0.35	0.333	18.6

* Calculated assuming equilibrium of the ^{238}U and ^{232}Th decay chains.

† Repeat measurements on three subsamples of 1a gave the same age within errors.

break up when placed in water. The other Scilly Isles sample, from St Mary's, (1f) had a similar texture but had large brown stains running through it. The sample from Barton in the New Forest area was also uniform but contained fine roots. Sample 1b from Sturt Pond in the New Forest area and 1e from the Lizard Peninsula, Cornwall, also contained roots and many small pockets of darker soil which are likely to have been washed in from above. The last two samples were initially considered to be unsuitable for TL dating because of likely contamination but were included for completeness.

Figure 1*a* shows typical glow curves for the Barton sample, showing natural TL and response to subsequent laboratory irradiations. To obtain the age of the sediment since its last exposure to sunlight, the natural TL signal needs to be compared with that induced in it by a known radiation dose. A $^{90}Sr/^{90}Y$ β source was used with method (c) of Wintle and Huntley[4] to separate the TL due to radiation exposure since burial from that which was not bleached by the original sunlight exposure (termed the residual component). Figure 1*b* shows the effect of additional doses of radiation on top of the natural TL for the same sample (on a different vertical scale): the lower temperature peaks are the more sensitive. Above 350 °C the residual component is responsible for >50% of the natural TL and this, combined with the low TL sensitivity in this region, makes it impossible to obtain the equivalent radiation dose (ED) in this region. The values of the ED obtained in the glow curve temperature region 260–320 °C are given in Table 1. In this region the value of the ED obtained does not vary with glow curve temperature. Integrating over this temperature range, the residual signal is equivalent to 40% of the total natural TL for samples 1b to 1f and 50% for 1a; ignoring it would have added about 3.5 krad to the ED of each sample. This is typical of residual signal obtained by the same technique for polymineral samples of ocean sediments having EDs in the range 3–26 krad (ref. 3). The TL signal from samples which have been irradiated after heating indicates that the TL at lower doses is not perfectly linear and an intercept correction (I) is made as in the case of the quartz inclusion TL dating method for pottery[5]. The error in the determination of the past radiation dose ($ED+I$) is $\sim\pm 7\%$ and is due mainly to the subtraction of the residual signal.

To obtain the age of the sediment, the radiation dose obtained from the TL measurements must be combined with the annual dose rate obtained from radioactivity analyses. The contribution from ^{40}K is calculated from the total K_2O content measured by atomic absorption spectrometry and for these samples the ionization due to ^{40}K is responsible for about a third of the natural TL. Apart from a 14 mrad yr^{-1} contribution due to cosmic radiation, the remaining radiation is due to the α, β and γ radiation from the decay of ^{238}U and ^{232}Th chains present in the sediment. Their activity is measured by a ZnS α scintillation counting technique[6]. Table 1 shows that the total α count rate is very similar for the four cleanest samples. The two which were observed to contain small pockets of darker soil had larger α counts, perhaps indicating recent contamination. (Sub-sampling of these specimens also gave a scatter of 10% whereas the cleaner samples were reproducible to within 5%. Clean water was dripped continuously through samples 1a and 1b for 10 days after measurement and on recounting no change was observed for 1a but a loss of 25% in total α count was observed for 1b.) If the U and Th decay chain contributions had been the same as for the other samples, the ages for 1b and 1e would have been 17,200 and 18,200 yr respectively.

The α efficiency factor, a, was measured for each sample and the annual dose rate calculated using the energy conversion factors of Bell[7] and assuming a mean water content of 20%. The saturation water content was measured for each sample in the laboratory. However, except for the two samples from the Scilly Isles, 1d and 1f, the samples formed a slurry and it was thought incorrect to try and estimate a mean annual water content from these values. The 20% value was found to be the mean *in situ* value for British brick-earth deposits. The water content is important because of the absorption of energy by the water which would otherwise be absorbed in the minerals to produce a higher TL signal. The age calculated is directly dependent on the per cent water content and this is the greatest source of error in dating these samples. Anomalous fading tests were carried out and no losses within the experimental reproducibility limit of ±4% were found in the 260–320 °C temperature range.

This is the first application of TL dating to late Devensian loess, although apparently successful TL dating of the quartz fraction of older loess has been reported in the Soviet Union[8]. The uncertainty in estimating the water content and the residual signal are the two major sources of error in the age and an overall error of ±20% should be applied to these dates; this makes it impossible to separate the six samples studied here but they are easily distinguished from older loesses which I am studying. TL dating can thus resolve the controversy concerning the presence of late and/or early Devensian loess in Britain.

I thank Dr J. A. Catt for samples and for stimulating discussion, Mrs M. Miller for potassium analyses and Dr M. J. Aitken for use of a ^{242}Cm α source. Financial support was provided by a NERC grant to Dr N. J. Shackleton.

Received 29 August; accepted 15 December 1980.

1. Catt, J. A. in *British Quaternary Studies, Recent Advances* (ed. Shotton, F. W.) 222–229 (Clarendon, Oxford, 1977).
2. Kerney, M. P. *Proc. geol. Ass.* **76**, 269–274 (1965).
3. Wintle, A. G. & Huntley, D. J. *Nature* **279**, 710–712 (1979).
4. Wintle, A. G. & Huntley, D. J. *Can. J. Earth Sci.* **17**, 348–360 (1980).
5. Fleming, S. J. *Thermoluminescence Techniques in Archaeology* (Clarendon, Oxford, 1979).
6. Aitken, M. J. *Physics and Archaeology* 2nd edn (Oxford University Press, 1974).
7. Bell, W. T. *Archaeometry* **19**, 99–100 (1977).
8. Shelkoplyas, V. N. *Geokhronologiya SSSR* Vol. 3 Ch. 2 (ed. Zubakova, V. A.) (1974).

Bermuda sea level during the last interglacial

Russel S. Harmon*, Lynton S. Land†, Richard M. Mitterer‡, Peter Garrett§, Henry P. Schwarcz‖ & Grahame J. Larson

* Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride G75 0QU, UK
† Department of Geological Sciences, University of Texas, Austin, Texas 78712
‡ Department of Geology, University of Texas at Dallas, Richardson, Texas 75080
§ Department of Geology, Colby College, Waterville, Maine 04901
‖ Department of Geology, McMaster University, Hamilton, Ontario, Canada L8S 4M1
Department of Geology, Michigan State University, East Lansing, Michigan 48824

Bermuda is a stable, mid-oceanic carbonate platform for which a particularly complete record of sea level fluctuation during the period ~135–70 kyr has been reconstructed from $^{230}\mathrm{Th}/^{234}\mathrm{U}$ dating of corals and speleothems, amino acid racemization dating of molluscs, and a reinterpretation of geological field relations. During the past 5 yr Bermuda has been the focus of an extensive geochronological programme which has included a detailed re-examination of the geology (to be reported elsewhere[1–3]). Here we focus only on the latest interglacial sequence. We use the stratigraphic nomenclature of Land *et al.*[4], but will present new age data and interpretations of geological relationships. This information, as it relates to the last interglacial period, is given in Tables 1 and 2 and summarized in Fig. 1. We show that during the last interglacial period, sea level was only above its present level for a few thousand years at about 125 kyr when it stood at +4 to +6m. At ~105 and ~85 kyr, sea level rose to between −15 and −20 m as constrained by ages for submerged speleothems and eolianites. At ~95 kyr sea level in Bermuda was at least −15 m below present level.

The Bermudas are a group of tectonically stable, Pleistocene-age islands in the mid-Atlantic (35°N, 65°W) consisting of intercalated and interfingering shallow-water marine calcarenites, eolianites and terra-rose palaeosols deposited on a stable volcanic pedestal. At least four complex, superimposed suites of marine and eolian carbonates occur in a general pattern of seaward accretion with the younger units lying closer to the coast. Within each suite there is an alternation of carbonate facies, with successive suites generally separated by palaeosol horizons. Caves formed by groundwater solution and collapse are common within the older, more lithified carbonate units and often extend below present sea level and are partially or totally filled with seawater. Extensive deposition of calcium carbonate speleothems (stalactites, stalagmites and flowstones) at elevations above and below present sea level is characteristic of the older caves.

Bretz[5] and Land *et al.*[4] established the conceptual framework within which Bermudian geology must be considered. They recognized that the carbonate formations are not simple stratigraphic units arranged in layer-cake fashion which can be separated solely on the basis of lithological or palaeontological criteria, and pointed out that an understanding of depositional mechanism is critical to the interpretation of Bermudian geology. Eolianites comprise over 90% of the exposed carbonate rock in Bermuda and represent lithified, transverse-lobate dune ridges[6]. These can be observed in some cases to merge laterally into marine carbonates and therefore must have been deposited during times of interglacial high sea stand. Thus, preserved littoral marine carbonates and patch reefs must have formed during interglacial periods at times when sea level was at or above its present level and eolianites can only have formed when the Bermuda platform (−20 m) was submerged to provide a source of sand for eolian shoreline dunes, but with sea level not necessarily as high as at present. During intervening glacial periods of depressed sea level, the Bermuda platform was subaerially exposed, and accumulation of terra-rose or regolith palaeosols and deposition of speleothems in caves at elevations below present sea level occurred coincident with vadose carbonate diagenesis. Thus, because the Bermuda platform is in an area of the world not significantly affected by either the effects of vertical displacement due to tectonic uplift or isostatic deformation through ice or water loading[7,8], Bermudian geology represents an approximate palaeosea level gauge, presumably recording absolute eustatic fluctuation of late Pleistocene sea level.

The last interglacial period in the marine record from ~135,000 to 75,000 yr BP is commonly assigned to Stage 5 of the marine oxygen isotopic record[9,10]. As Fig. 1 shows, the Bermuda platform was probably submerged for most of this time. Before Stage 5, during Shore Hills time, sea level was below the edge of the Bermuda platform, then rose rapidly, passing −13 m at ~134 kyr, to the Devonshire +4 to +6 m high stand at 125 kyr. Here it produced a wave-cut notch in the older Belmont Formation, deposited the Devonshire marine calcarenites and flooded what are at present the air-filled portions of many caves resulting in the development of aragonitic surface coatings on actively forming stalagmites and stalactites up to elevations of +3 to +4 m. Sixteen $^{230}\mathrm{Th}/^{234}\mathrm{U}$ dates on Devonshire corals from sites around Bermuda where stratigraphic relationships are clear range from 134 ± 8 to 118 ± 6 kyr with a mean of 125 ± 4 kyr (Table 2), thus correlating the Devonshire Formation with the Rendezvous Hill terrace in Barbados[11,12], reef complex VII in New Guinea[13] and the Wiamanalo shoreline in Hawaii[14]. Sea level stood at this level for only a short period before quickly falling off the Bermuda platform as indicated by initiation of speleothem deposition at elevations below present sea level by ~120 kyr. This regression is reflected in the geological record by the development of the Harrington Formation which gives both U-series and amino acid racemization ages of ~120 kyr (Table 2).

The remainder of Stage 5 time in Bermuda is a long and complex episode of dune building and soil deposition resulting from periodic submergence and exposure of the Bermuda platform but with sea levels always below present level. The amino acid racemization ages of ~105 and ~86 kyr for two post-Devonshire eolianites (Table 2) record the timing of these transgressions. The dates correlate well with isotope Stages 5c

Table 1 New $^{230}\mathrm{Th}/^{234}\mathrm{U}$ ages for Bermuda stalagmites

Sample no.	Depth (m)	U conc. (p.p.m.)	$(^{234}\mathrm{U}/^{238}\mathrm{U})$	$(^{230}\mathrm{Th}/^{234}\mathrm{Th})$	$(^{230}\mathrm{Th}/^{232}\mathrm{Th})$	Age (kyr BP)
75008:02(t)	−3	0.08	1.04 ± 0.04	0.47 ± 0.03		68 ± 6
:01(b)	−3	0.07	1.11 ± 0.03	0.61 ± 0.03	>200	97 ± 9
77520:01(t)	−15	0.14	1.17 ± 0.03	0.09 ± 0.02	30	10 ± 2
:03(m)	−15	0.09	1.05 ± 0.04	0.30 ± 0.05	>200	39 ± 7
:02(b)	−15	0.08	1.60 ± 0.02	0.67 ± 0.03	>200	111 ± 9
77522:01(t)	−13	0.08	1.53 ± 0.02	0.75 ± 0.03	110	134 ± 11

t, top; m, middle; b, base.

0028–0836/81/060481—03$01.00

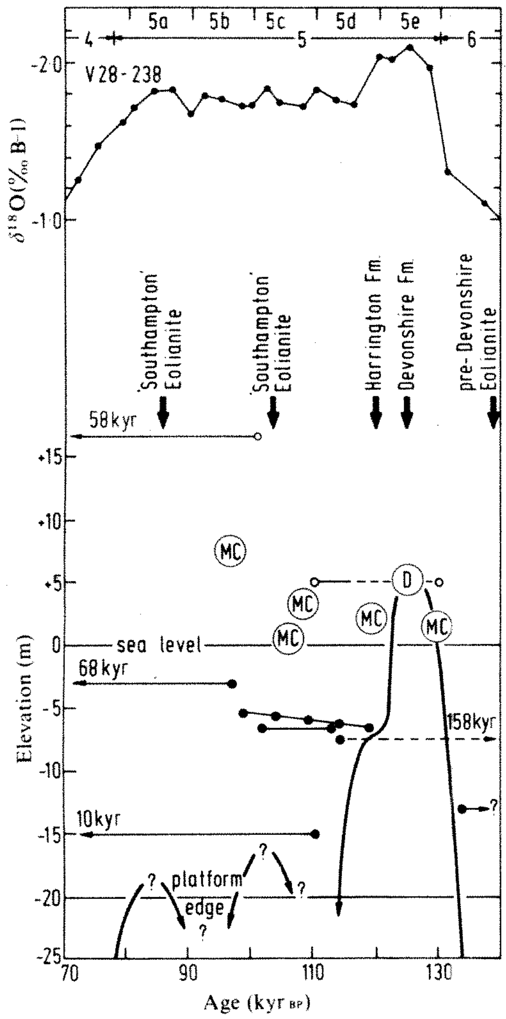


Fig. 1 Palaeosea level curve for Bermuda during the last interglacial. ^{230}Th/^{234}U age determinations: ●, submerged speleothems; ○, aerially exposed speleothems. Solid lines connecting speleothem age determinations indicate continuous deposition over the indicated interval; dashed lines signify interrupted growth. Horizontal arrows denote speleothem deposition outside the 70–140 kyr period of interest here. The circle labelled 'D' indicates the Devonshire last interglacial maximum high sea stand; those labelled 'MC' indicate time–elevation relationships for south-shore storm deposits (conglomerates). Vertical arrows show amino acid racemization ages for molluscs from the stratigraphic units indicated calibrated to D. The sea-level curve is based entirely on Bermudian geology and chronology. Correlation between the Bermuda sea level and the marine oxygen isotope record[10] is based on absolute dating. The isotope stages are numbered after Emiliani[9].

and 5a (ref. 10), as well as the Ventnor and Worthing terraces in Barbados[11,12] and reef complexes VI and V in New Guinea[13]. This is the first time to our knowledge that these events have been documented in a tectonically stable area. The deposition of stalagmites at elevations down to −15 m indicates that sea level was below this level from ~110 kyr to the end of Stage 5. This is consistent with the fact that no Pembroke or Southampton marine facies or strandlines have been observed, as they must be below present sea level.

The two most persistent problems of Bermudian geology have been the patchy marine conglomerates which occur along the south shore at elevations between sea level and +11 m, and a complex of evenly bedded, seaward-dipping sands which are present along the north shore at elevations up to about +20 m. Land *et al.*[4] assigned the conglomerates to one lithological unit, the Spencer's Point Formation, which they interpreted to be a single post-Pembroke marine transgression. These conglomerates and the north-shore sands were then correlated because the latter were considered marine in character near sea level and can be traced inland tomverlap what was considered to be Pembroke eolianite. However, ^{230}Th/^{234}U dating of corals from seven of these conglomerates (Table 2) has yielded ages correlative with each transgressive phase of Stage 5, plus one modern age. Close examination of these conglomerates in the field reveals each to be an isolated, chaotic deposit emplaced as a plaster or cliff pocket infilling. Thus, we reinterpret them to be deposits thrown up along the dynamic south-shore coastline from storms during times of platform submergence. This result has two important consequences. First, it emphasizes that lithological units as mapped in the field (that is rock-stratigraphic units) are not necessarily time-stratigraphic units. Second, it removes the correlation between the south-shore conglomerates and north-shore sands.

The north-shore sands we now assign to the Devonshire Formation on the basis of amino acid racemization ages which are the first absolute age data for these deposits. The sea stand responsible for the deposition of the north-shore sands, 'perhaps as high as 20 m (Blackwatch Pass)'[4], has always been based on interpretation of depositional environment of the uniform seaward-dipping section. Some geologists might claim that all the units are marine, others have argued that they are all eolian

Table 2 Stratigraphic column for the last interglacial sequence of Bermuda[*]

Southampton Formation
  Poorly consolidated dunes with unconsolidated zones rich in land snails. Age >30 kyr by radiocarbon[*][†] and ~86 and ~105 kyr for molluscs from two different outcrops by amino[‡] acid racemization.

St George's Soil
  Red residual soil of moderate development and only rare palmetto stump casts.

Spencer's Point Formation
  Discontinuous patches of marine conglomerate along the south shore and harbour islands plastered in cliff pockets and recesses or on large breakdown blocks at elevations of between +1 to +11 m; previously correlated with seaward-dipping sands along the north shore[4]. Age: corals from various conglomerates have given ^{230}Th/^{234}U ages of 131±7, 124±8; 121±6, 118±6, 118±11, 108±6, 105±6, 97±6[‡], 85±6, 83±5, and <2 kyr.

Pembroke Formation
  Extensive, island-wide eolianites with landward-dipping foreset beds which commonly extend down to and below present sea level.

Harrington Formation
  A yellowish to reddish unconsolidated layer on the Devonshire or Belmont Formations which contains abundant land snails near the top and grades downwards into coarse marine beds near the base. Age: ~120 kyr by amino acid racemization for molluscs, and 121±9 by ^{230}Th/^{234}U.

Devonshire Formation
  Intertidal marine conglomerates and calcarenites at +2 to +6 m which contain abundant Belmont clasts and a characteristic fauna. Age: corals from this unit have given ^{230}Th/^{234}U ages of 118±6[‡], 124±6[‡], 124±8[‡], 124±6[‡], 124±12, 124±12, 125±6, 127±9, 127±9, 127±6, and 134±8 kyr.

'Unnamed Eolianite'
  Seaward-dipping eolianite underlying the Devonshire marine beds on the north and south shores. Age: ~140 kyr by amino acid racemization for molluses.

Shore Hills Soil
  Well developed red palaeosol of island-wide extent with abundant palmetto stump casts.

[*] After ref. 4.
[†] Data from ref. 18.
[‡] Data from ref. 1.

Table 3 Comparison of sea level elevations during the last interglacial period

Locality	Time (kyr)	Sea level (m)	Ref.
Bermuda	125	+4 to +6	16, this study
	118	>-20	
	105	-15 to -20	
	95	>-20	
	85	-15 to -20	
New Guinea	120	+8	13
	110	-40	
	105	-14	
	82	-14	
Barbados	125	+5	11
	105	-43±20	
	82	-42±20	
Marine oxygen isotope record*	125	+11	10
	116	-19	
	110	-	
	105	-17	
	95	-18	
	84	-11	

* For core V28-238 assuming $-0.1‰$ change in $\delta^{18}O$ of foraminifera calcite per 10 m rise in sea level.

backset beds[3]. We have found no definite evidence but have recognized apparent beach cusps as high as +3 m above present sea level, and conclude that higher beds must be due to reworking during storms (hurricanes) during Devonshire time. The underlying eolianites, which give an amino acid age of ~140 kyr, thus represent an early phase of the Devonshire transgression, during which time the extensive North Lagoon could have supplied sand and stone[5].

Because the Bermudian palaeosea level record for the last interglacial presented here (Fig. 1) is particularly complete, it has important implications for the idea of glacial surge. Hollin[15] has recently argued that a catastrophic surge of the East Antarctic ice sheet caused a rise in worldwide sea level to about +16 m at ~95 kyr. Although such an event should have left some evidence of its occurrence in the geological record of island and coastal areas, it can be argued that it might not be reflected in either the marine oxygen isotope or tectonic island sea level records because both of these are relatively low-resolution measures of global ice volume integrated over a time scale of a few thousand years. Hollin[15] recognized that his hypothesis must be tested in an area where a highly resolved sea level record for the last interglacial is preserved. Bermuda is one such locality, and thus the palaeosea level record presented here is a crucial test for such an ice surge.

Although Bermuda is cited by Hollin[15] to support a +16-m high sea stand at ~95 kyr, the reappraisal of geological relationships resulting from our chronological and field studies has revised the tenuous correlation on which a possible +20 m post-Devonshire high sea stand was based. Figure 1 shows that sea level in Bermuda was only once above its present level during Stage 5; that is during Devonshire time at 125 kyr when it stood at +4 to +6 m. Otherwise, from ~135 to 70 kyr sea level was below ~-15 m. There is no geological evidence to the contrary. There is no indication of a former freshwater table at elevations above present sea level, except the +2 m feature relating to the penultimate high sea stand which deposited the Belmont Formation between ~230 and 200 kyr (ref. 1). Neither have we observed any overgrowths nor surface coatings of marine aragonite and serpulid worm shells on any speleothems above +3 to +4 m in the more than 20 caves investigated nor found any interior aragonite layers in speleothems younger than 120 kyr. Such features are present to some extent on the surface of all speleothems presently submerged in seawater.

In Crystal Cave, stalactites at +2 to +4 m above sea level typically contain one interior layer of such marine deposition which can be easily observed because the lower portions of these speleothems have been broken by human activity to reveal a cross-section through these deposits. For one such stalactite previously studied, the aragonitic layer only extended upwards inside the stalactite to an elevation of +3 to +4 m, above which calcite crystal growth was continuous across the axis of the entire stalactite. $^{230}\mathrm{Th}/^{234}\mathrm{U}$ dating gave an age of 130±10 kyr for the calcite layer just inside the marine layer and 110±14 kyr for the outer calcite layer thus independently documenting the absolute elevation of the Devonshire high sea stand at 125 kyr (Fig. 1) (ref. 16). Certainly one would expect to observe two such sediment layers in these stalactites and marine overgrowths and surface coatings on subaerial travertines and flowstones as well as evidence for an elevated freshwater table between +4 and +16 m if there had been a post-Devonshire sea level rise to +16 m. Also, neither the extensive Devonshire marine calcarenites and relatively unlithified Pembroke or Southampton eolianites, nor the wholly unconsolidated, Harrington Formation have been reworked to any extent. All these units are exposed along the south shore of Bermuda at or near sea level, not more than 1–2 km from the platform edge, at present a very dynamic, high-energy environment. It, therefore, seems most unlikely that a +16 m rise in sea level over only a few hundred or thousand years, as required by the surge hypothesis[15], could have happened without some reworking, eroding, and re-depositing of these units. The discontinuous marine conglomerates cannot be surge-derived unless the coincidence of their ages with the well recognized isotope Stage 5e, 5c and 5a high sea stands is fortuitous, with surge events occurring within each of these transgressions, and demonstration that deposition of stalagmites below present sea level could have been interrupted by a marine transgression and then resumed without leaving behind physical or petrographical evidence to this effect.

Finally, the Bermuda palaeosea level record presented here is in basic agreement with both the marine oxygen isotope[10] as well as the New Guinea[13] and Barbados[11,12] sea level records (Table 3). Although there are some internal inconsistencies among these various ice-volume records which reflect both geological and chronological uncertainties, the overall agreement is striking. However, although casting serious doubt on Hollin's[15] postulated surge of the East Antarctic ice sheet at 95 kyr, these results do not question either the theory or reality of the general concept as discussed by Aharon *et al.*[17]. The Devonshire high sea stand may in fact represent a surge event at 125 kyr superimposed on the general submergence of the Bermuda platform during the last interglacial period.

The $^{230}\mathrm{Th}/^{234}\mathrm{U}$ coral and speleothem dating programme and palaeosea level work in Bermuda has been supported by a grant from the NSF (EAR 77-23026) and a Wilkinson Fellowship from the Bermuda Biological Station. The manuscript was typed by Miss J. McLean and illustrations drawn by J. Bennett.

Received 3 July; accepted 18 November, 1980.

1. Harmon, R. S. *et al. Palaeogr. Palaeoclimatol. Palaeoecol.* (in the press).
2. Mitterer, R. M., Land, L. S., Harmon, R. S. & Schwarz, H. P. (in preparation).
3. Garrett, P. & Vacher, H. L. (in preparation).
4. Land, L. S., Mackenzie, F. T. & Gould, S. J. *Bull. geol. Soc. Am.* **78**, 993–1006 (1973).
5. Bretz, J. H. *Bull. geol. Soc. Am.* **71**, 1729–1754 (1960).
6. Mackenzie, F. T. *Sedimentology* **3**, 51–64 (1964).
7. Walcott, T. I. *Quat. Res.* **2**, 1–14 (1973).
8. Bloom, A. L. in *Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 355–379 (Yale University Press, 1971).
9. Emiliani, C. *J. Geol.* **63**, 538–578 (1955).
10. Shackleton, N. J. & Opdyke, N. D. *Geol. Soc. Am. Mem.* **14** 449–464 (1976).
11. Mesolella, K. J., Matthews, R. K., Broecker, W. S. & Thurber, D. L. *J. Geol.* 250–274 (1969).
12. Bender, M. L. *et al. Bull. geol. Soc. Am.* **90**, 577–597 (1979).
13. Bloom, A. L., Broecker, W. S., Chappell, J. A., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 184–205 (1974).
14. Ku, T. L., Kimmel, M. A., Easton, W. H. & O'Neil, T. J. *Science* **183**, 959–962 (1974).
15. Hollin, J. T. *Nature* **281**, 629-633 (1980).
16. Harmon, R. S., Schwarz, H. P. & Ford, D. C. *Quat. Res.* **9**, 205–218 (1978).
17. Aharon, P., Chappell, J. & Compston, W. *Nature* **283**, 649–651 (1980).
18. Velastro, S., Davis, E. M. & Varela, A. G. *Radiocarbon* **14**, 461–485 (1972).

Low $^{15}N/^{14}N$ in hydrothermal vent animals: ecological implications

Greg H. Rau

Institute of Geophysics and Planetary Physics,
University of California, Los Angeles, California 90024

Low-trophic-level animal biomass from the Galápagos and 21 °N hydrothermal vents was found to have $^{15}N/^{14}N$ ratios which are lower than those previously reported for organic nitrogen[1–6], ammonium[1,2,5], and nitrate[1,2,7,8] in deep Pacific Ocean water. These nitrogen isotope ratios of vent animals are, however, near those of deep-oceanic N_2 (refs 1, 2, 7, 9) and resemble those of marine biota associated with N_2 fixation[2,10]. These observations suggest that organic nitrogen of nutritional importance to vent animals is initially synthesized within the vent environment and this synthesis may be preceded by N_2 fixation.

It has been argued[11–19] that local bacterial chemo-autotrophy is the primary source of biomass and energy for the dense invertebrate communities found at Pacific Ocean hydrothermal vents. Part of the evidence for vent primary production is the difference between the $^{13}C/^{12}C$ of vent animals and the $^{13}C/^{12}C$ of animals in the overlying pelagic environment[13,19]. It was reasoned that measurement of vent animal $^{15}N/^{14}N$ would provide further clues as to the base and the structure of these vent food webs.

A variety of animal tissues from hydrothermal vents were analysed including: trophosome and vestimentum from *Riftia pachyptila* Jones (a vestimentiferan worm), mantle tissue from *Calyptogena magnifica* Boss and Turner (vesicomyid clams), and crab muscle from *Bythograea thermydron* Williams (brachyuran crabs). These animals had been originally taken from either the Galápagos or the 21° N hydrothermal vent sites through the submersible Alvin, as indicated in Fig. 1. These tissues were dried at 60 °C and ground, and the nitrogen present in 5–10-mg replicate subsamples of these tissues was completely converted to N_2 using methods described by Stump and Frazer[20]. The $^{15}N/^{14}N$ relative to air (Fig. 1 legend) of these N_2 gas samples was determined by David Winter using a Varian MAT 250 ratio mass spectrometer.

Figure 1 shows vent animal $\delta^{15}N$ ranges from +1.8 to +9.8‰, with the values generally increasing as a function of the animal's assumed trophic level. This animal $\delta^{15}N$ increase with trophic status corroborates observations[2,4,10,21–25] that food chain/heterotrophic processes tend to increase ^{15}N relative abundance in the resultant animal biomass. Consequently, if vent animals were utilizing sedimentary organic nitrogen that had been produced and biologically processed on its journey through the more than 2 km of water overlying the vents, one would expect the $\delta^{15}N$ of vent animals: first, to exceed the $\delta^{15}N$ of plankton encountered in the surface layers of the Pacific Ocean, commonly 5–10‰ (refs 1, 2, 5, 10, 26), and second, to equal, if not exceed, the $\delta^{15}N$ of deep-ocean sedimentary organic nitrogen, 5–13‰ (refs 1–6). The relatively low $\delta^{15}N$ of assumed primary producer–consumer tissues from the vents, 1.8–4.9‰ (Fig. 1), however, suggests that the organic nitrogen present in these tissues has undergone relatively little biological cycling and is, therefore, of local, rather than pelagic–sedimentary origin.

Additionally, the striking similarity between the $\delta^{15}N$ of low-trophic level vent animal tissues and the $\delta^{15}N$ of marine organisms associated with N_2 fixation (Fig. 1) suggests that the production of the organic nitrogen found in vent animals is preceded by N_2 fixation. Both the concentration and the low $\delta^{15}N$ of dissolved N_2 are relatively constant throughout the ocean[1,7,9], and little nitrogen isotopic fractionation is known to occur during fixation of N_2 and subsequent organic nitrogen synthesis[2,10,27–31]. The yet-to-be measured $\delta^{15}N$ of the NH_4^+ and the NO_3^- present[32] in the vicinity of the vents may, however, prove these ions as likely as N_2 to be precursors to vent animal nitrogen. Large nitrogen isotope fractionations incurred during nitrate or ammonium assimilation and biomass production by some microbes and by some autotrophs in certain conditions[2,28,31,33–35] may also be factors causing vent organisms to be relatively depleted in ^{15}N.

Whatever the synthesis pathways and their inorganic starting compounds, the existing isotopic evidence indicates that the organic nitrogen as well as the organic carbon[13,19] found in vent animal biomass are largely produced by non-pelagic and non-photosynthetic processes.

This research was made possible through the generosity of Mr Bud Smithey and Drs Robert Hessler, George Somero, Horst Felbeck, James Childress and Ian Kaplan. Support was provided in part by a joint US Bureau of Land Management/Department of Energy contract No. EY-76-3-03-0034

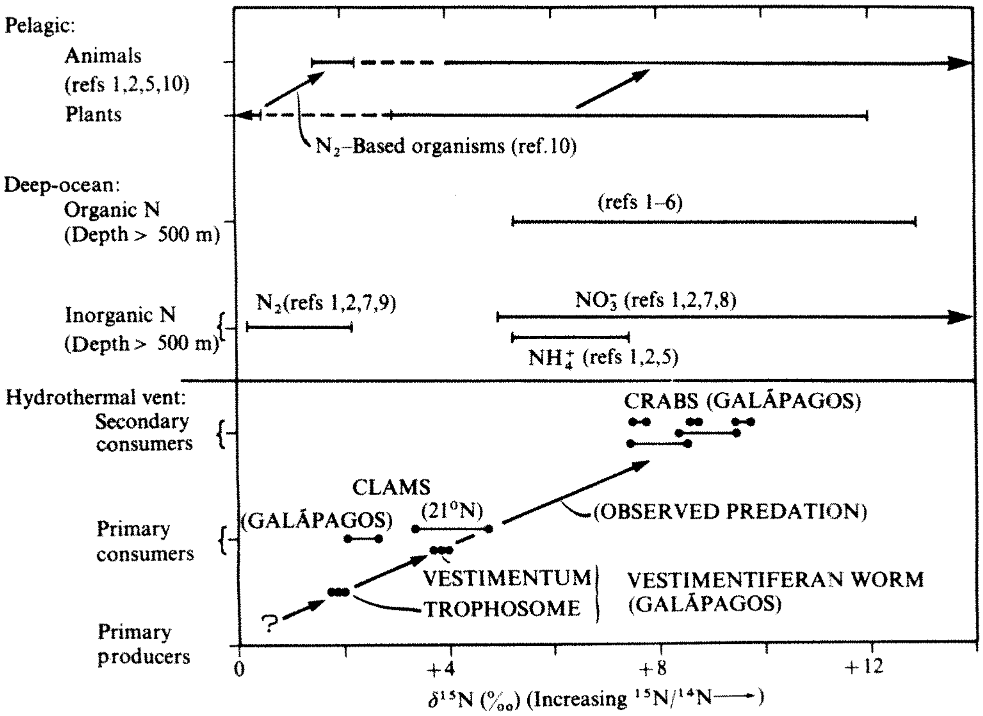


Fig. 1 The $\delta^{15}N$ of various Pacific Ocean nitrogen pools where:

$$\delta^{15}N = \left\{\left[\frac{(^{15}N/^{14}N)_{sample}}{(^{15}N/^{14}N)_{air}}\right] - 1\right\} \times 10^3‰$$

The vent sites, Galápagos (0.8° N, 86.1° W) and 21° N (20.9° N, 109.1° W), are located at ocean depths of 2.5 and 2.6 km, respectively. The assumed trophic levels and transfers (diagonal arrows) of hydrothermal vent animals/tissues are based on knowledge of related species, and on observations by Drs R. R. Hessler, G. N. Somero, H. Felbeck and J. J. Childress. The trophosome of the vestimentiferan worm is at least partly, if not largely, composed of autotrophic tissues[16–18]. Replicate analyses of a single individual's tissue are connected by a horizontal line.

0028–0836/81/060484—02$01.00

with UCLA, as well as NSF grants OCE-78-08852 (to Childress), OCE 78-08853 (to Somero), and OCE 78-10460 (to Hessler). This is contribution no. 23 of the Galápagos Rift Biology Expedition supported by the NSF.

Received 11 August; accepted 7 November 1980.

1. Miyake, Y. & Wada, E. *Rec. Oceanogr. Works Jap.* **9,** 137–152 (1967).
2. Wada, E., Kadonoga, T. & Matsuo, S. *Geochem. J.* **9,** 139–148 (1975).
3. Peters, K. E., Sweeney, R. E. & Kaplan, I. R. *Limnol. Oceanogr.* **23,** 598–604 (1978).
4. Saino, T. & Hattori, A. *Nature* **283,** 752–754 (1980).
5. Sweeney, R. E. & Kaplan, I. R. *Mar. Chem.* **9,** 81–94 (1980).
6. Venkatesan, M. I., Brenner, S., Ruth, E., Bonilla, J. & Kaplan, I. R. *Geochim. cosmochim. Acta* **44,** 789–802 (1980).
7. Cline, J. D. & Kaplan, I. R. *Mar. Chem.* **3,** 271–299 (1975).
8. Liu, K. K. thesis, UCLA (1979).
9. Benson, B. B. & Parker, P. D. M. *Deep-Sea Res.* **7,** 237–253 (1961).
10. Wada, E. & Hattori, A. *Geochim. cosmochim. Acta* **40,** 249–251 (1976).
11. Corliss, J. B. *et al. Science* **203,** 1073–1083 (1979).
12. Jannasch, H. W. & Wirsen, C. O. *Bioscience* **29,** 592–598 (1979).
13. Rau, G. H. & Hedges, J. I. *Science* **203,** 648–649 (1979).
14. Karl, D. M., Wirsen, C. O. & Jannasch, H. W. *Science* **207,** 1345–1347 (1980).
15. Hiatt, B. *Mosaic* **11,** 15–21 (1980).
16. Cavanaugh, C., Gardiner, S. L., Jones, M. L., Jannasch, H. W. & Waterbury, J. B. *Science* (submitted).
17. Felbeck, H. *Science* (submitted).
18. Jones, M. L. *Science* (submitted).
19. Rau, G. H. *Science* (submitted).
20. Stump, R. K. & Frazer, J. W. *Nucl. Sci. Abstr.* **28,** 746 (1973).
21. Gaebler, O. H., Choitz, H. C., Vitti, T. G. & Vukmirovich, R. *Can. J. Biochem. Physiol.* **41,** 1089–1097 (1963).
22. Gaebler, O. H., Vitti, T. G. & Vukmirovich, R. *Can. J. Biochem.* **44,** 1249–1257 (1966).
23. Pang, P. C. & Nriagu, J. O. *Geochim. cosmochim. Acta* **41,** 811–814 (1977).
24. Steele, K. W. & Daniel, R. M. *J. Agric. Sci. Camb.* **90,** 7–9 (1978).
25. DeNiro, M. J. & Epstein, S. *Geochim. Cosmochim. Acta* (in the press).
26. Sweeney, R. E., Liu, K. K. & Kaplan, I. R. in *Stable Isotopes in the Earth Sciences* (ed. Robinson, B. W.) 9–26 (DSIR, Wellington, 1978).
27. Hoering, T. C. & Ford, H. T. *J. Am. chem. Soc.* **82,** 376–378 (1960).
28. Delwiche, C. C. & Steyn, P. L. *Envir. Sci. Technol.* **4,** 929–935 (1970).
29. Delwiche, C. C., Zinke, P. J., Johnson, C. M. & Virginia, R. H. *Bot. Gaz.* **140**, Suppl. S65–S69 (1979).
30. Shearer, G. & Kohl, D. H. in *Recent Developments in Mass Spectrometry in Biochemistry and Medicine* Vol. 1 (ed. Frigerio, A.) 605–622 (Plenum, New York, 1978).
31. Kohl, D. H. & Shearer, G. *Plant Physiol.* **66,** 51–60 (1980).
32. Edmonds, J. M. *et al. Earth planet. Sci. Lett.* **46,** 19–36 (1979).
33. Wada, E. & Hattori, A. *Geomicrobiol. J.* **1,** 85–101 (1978).
34. Kohl, D. H. & Shearer, G. in *Recent Developments in Mass Spectrometry in Biochemistry and Medicine* Vol. 1 (ed. Frigerio, A.) 623–640 (Plenum, New York, 1978).
35. Wada, E. in *Isotope Marine Chemistry* (eds Goldberg, E. D., Horibe, Y. & Saruhashi, K.) 375–398 (Uchida Rokakuho, Tokyo, 1980).

Fourier analysis of test shape of planktonic foraminifera

Nancy Healy-Williams & Douglas F. Williams

Belle W. Baruch Institute for Marine Biology and Coastal Research
and Department of Geology, University of South Carolina,
Columbia, South Carolina 29208

The exact characterization of test shape is important in morphological studies of foraminifera and other marine microfossils. Most biometric studies of microfossils, however, have only measured test diameter or other linear components[1]. The tests of planktonic foraminifera exhibit a wide range of morphological traits, some of which are related to environmental conditions and others to the phylogenetic history of the taxon[1,2]. Mean test size or linear measurements of chambers and/or whole tests of fossil foraminifera can be used for palaeoecological interpretations from the morphological variations[3–10]. We report here that Fourier shape analysis could be an objective and unambiguous biometric method for determining the components which comprise the shape of foraminifera. Fourier shape analysis has been successfully applied to fossil ostracods[11,12], fossil bryozoans[13,14] and miospores[15]. In the present study, Fourier series shape analysis was performed on the planktonic foraminifera, *Globorotalia truncatulinoides*, from the southern Indian Ocean, to differentiate palaeoenvironmentally significant trends in the test shape of this species. Kennett[16] and Takayanagi *et al.*[17] have used linear test measurements to describe latitudinal gradients in the test morphology of *G. truncatulinoides*. Highly conical forms dominate in tropical waters, whereas more compressed, biconvex forms dominate in cool water.

Table 1 Core top locations used in the Fourier analysis of *Globorotalia truncatulinoides*

Core	Lat (°S)	Long (°E)
V20–175	22° 18′	68° 00′
RC11–145	25° 29′	110° 01′
E48–11A	29° 40′	97° 32′
E48–46A	31° 09′	83° 46′
E48–32A	32° 48′	83° 52′
E48–31A	34° 49′	84° 08′
E48–18	35° 26′	91° 53′
E48–28A	38° 33′	79° 55′
E48–23	39° 31′	83° 43′
E48–03	41° 01′	100° 00′
E50–03	42° 01′	100° 00′
E49–46A	44° 04′	100° 01′
E45–79	45° 03′	114° 22′
E45–77A	46° 26′	114° 25′
E49–42A	47° 20′	100° 00′

Closed-form Fourier series analysis is a mathematical expression of cosines used to describe any shape or any two-dimensional shape outline to a specified accuracy[18]. Fourier series both provides exact characterization of shape and separates the components of the whole form, enabling us to see how each component (harmonic) affects the total shape. Each harmonic is an independent and uncorrelated measure of a specific geometric contribution to the total shape[14,19]. The shape of each specimen is represented by a series of harmonics of the form: $Rn \cos(n\theta - \phi n)$, where Rn represents the amplitude of the nth harmonic, and the remainder concerns the phase angle of that harmonic. The set of harmonic amplitudes constitutes the 'harmonic amplitude spectrum' (Fig. 1*a*). The Fourier series is estimated for each planktonic foraminiferal projection from analysis of the coordinates of peripheral points with the origin at the centre of gravity of the foraminifera (see ref. 18). Each harmonic amplitude represents the relative contribution to the empirical shape of a characteristic shape component. For example, the amplitude of the second harmonic represents the contribution to the empirical shape of a 'figure eight' and is thus a measure of elongation. The third harmonic represents a trefoil, the fourth a quadrefoil, and so on. In general, the nth harmonic amplitude represents the shape contribution of an n-leaved clover[18]. The observed shape is thus partitioned into a series where gross shape (for example, elongation, and triangularity) is measured by the harmonic amplitudes of the lower harmonic orders, and increasingly fine-scaled surface sculpture is measured at higher orders.

Fourier analysis in closed form, given enough terms, has enough information to describe any shape completely. This is demonstrated in Fig. 2 by computing the Fourier series of amplitudes and phase angles for a given shape, and then using only the information contained in the Fourier series to reconstruct the shape. For example, in Fig. 2, composite shapes are formed for a specimen of *G. truncatulinoides* by the successive addition of the first 20 terms in the computed Fourier series until the original shape is accurately reproduced.

Table 2 χ^2 Contingency table*

Harmonic no.	χ^2 Results
2	15.3†
3	49.9†
4	25.2†
5	4.43
6	30.06†
7	6.5
8	10.69
9	10.68
10	9.63

* Calculations were performed on two end member samples: E48–11A and E49–42A (Table 1).

† Values that exceed critical χ^2 value (0.1) of 13.3.

0028–0836/81/060485—03$01.00

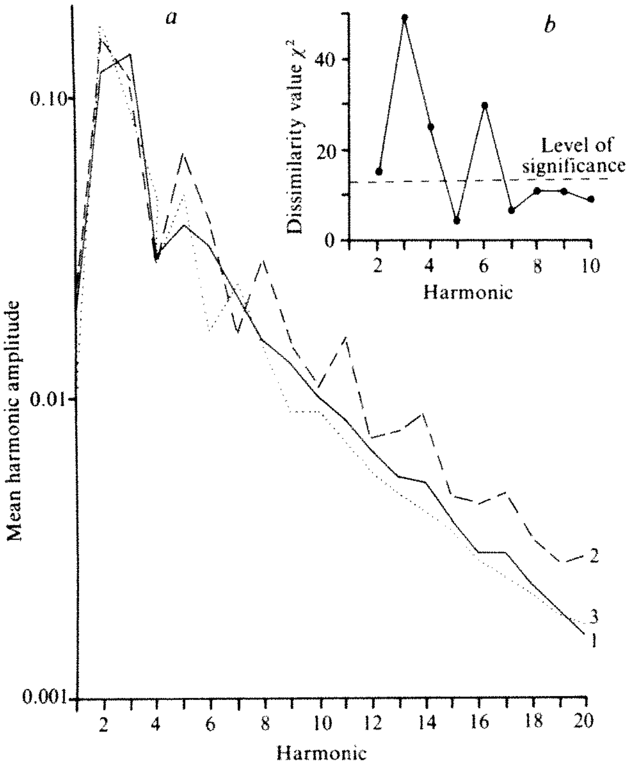


Fig. 1 *a*, Mean harmonic amplitude spectra of *G. truncatulinoides* from (1) V20–175, (2) E48–28A and (3) E49–42A (Table 1); *b*, plot of χ^2 values indicating the significance of harmonics 2, 3, 4 and 6.

Table 3 Correlation coefficients from linear regression analyses

	Harmonic 2 (elongation)	Harmonic 3 (conical shape)	Surface temperature (°C)	Coiling % sinistral	Lat
Harmonic 2	1.00	−0.46	−0.70	0.87	0.76
Harmonic 3	−0.46	1.00	0.77	−0.42	−0.73
Temperature (°C)	−0.70	0.77	1.00	−0.76	−0.99
Coiling	0.87	−0.42	−0.76	1.00	0.81
Lat	0.76	−0.73	−0.99	0.81	1.00

Values are results from Fourier shape analysis of *G. truncatulinoides*.

The specimens of *G. truncatulinoides* used in this Fourier biometric study were removed from the surface sediment of 11 trigger cores and four piston cores from the southern Indian Ocean (lat 22–47° S) (Table 1). The principal oceanographic features of the surface waters in this region are the subtropical convergence (38–40° S) separating the subtropical and sub-Antarctic waters, and the polar front (48–50° S) separating sub-Antarctic and Antarctic waters. Both fronts are defined by strong salinity and temperature gradients[20]. Only samples which contained well preserved Recent or late Holocene foraminifera were chosen[21,22]. An aliquot was removed from each sample using an FB Dial microsplitter and ~35 foraminiferal specimens were selected from the 300–450 μm size fraction. A separate experiment was performed to monitor the effects of size on test shape[23]. The coiling direction of the assemblages varied from 100 to 6% sinistrally coiled. Fourier analysis was performed on the side view of all specimens. A normalization process in the Fourier series means that differences in size do not affect shape comparisons. The periphery of each foraminifera specimen was digitized as 48 equiangular coordinates and then converted to a Fourier series containing 20 harmonics. As seen in Fig. 2, the first few harmonics describe the gross shape while the higher-order harmonics refine the overall test shape. The result is a harmonic amplitude spectrum for each foraminiferal specimen with changes in shape reflected by changes in amplitude values. The amplitude spectrum for a sample comprises the primary data used for analysis (Fig. 1*a*). Each foraminifera in a sample contributes an amplitude value for every term of the series, that is the 20 terms calculated produce an $n \times 20$ matrix (n = number of foraminifera). Thus, for each term of the series, histograms (Fig. 3) can be constructed by plotting amplitude values (Fig. 1*a*) as a function of frequency of occurrence. The histograms are called shape frequency distributions. Each foraminifera can be described by 20 harmonic amplitude values, so each sample can be described by 20 shape frequency distributions (Fig. 3).

Thus, with the *G. truncatulinoides* data in this form, samples can be statistically compared and the specific harmonics responsible for shape differences can be determined. A χ^2 table was constructed for the first 10 harmonics as these harmonics capture the significant components of overall shape in this species (Fig. 1*b*, Table 2). According to Benzecri[24], χ^2 values can be used as indicators of dissimilarity. Plots of χ^2 values indicate that the 2nd, 3rd, 4th and 6th harmonics are statistically significant in describing the shape of this species (Fig. 1*b*). Because the harmonic amplitude spectra indicated that harmonics 2 and 3 contain the highest degrees of shape information (Fig. 1*a*), these are used in the following analysis.

Shape frequency diagrams for harmonics 2 (elongation) and harmonic 3 (triangularity or conical shape) show the amplitude distributions in all 15 samples ranked according to latitude (Fig. 3). The conical shape of *G. truncatulinoides* (harmonic 3) increases from the subpolar to subtropical water masses, in agreement with earlier biometric studies of this species[16,17]. The frequency distributions of harmonic 2 (elongation), however, do not show a clear relationship with latitude. Instead, elongation seems related to the coiling direction of the populations (Fig. 3). In populations dominated by sinistrally coiled individuals (80–100% sinistral) south of 31° S, the mean amplitude values do not change systematically and show no variation in the elongation of *G. truncatulinoides*. In populations with significant numbers of dextrally coiled individuals (42–94% dextral) (30–22° S), *G. truncatulinoides* becomes less elongated. This relationship has never before been observed in this species.

Linear regression analyses were performed on harmonics 2 and 3 in relation to coiling ratio and various environmental parameters (Table 3). Harmonic 2 is strongly correlated with coiling direction ($r = +0.87$). Elongation is also correlated with latitude and surface-water temperature but not at the level of significance as with coiling-direction preference. Harmonic 3, on the other hand, has its highest correlation with surface-water temperature ($r = 0.77$) and a poor correlation with coiling ($r = 0.42$). We interpret the correlation between triangularity and water temperature to be indicative of ecophenotypic variation, that is, environmental-induced variation. The elongation (harmonic 2) of *G. truncatulinoides* may reflect more phenotypic

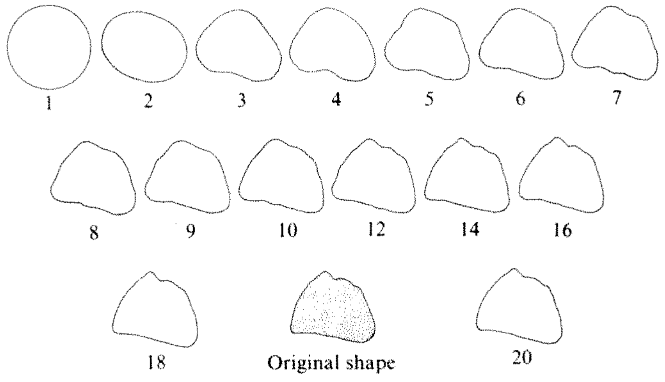


Fig. 2 Reproduction of the shape of *G. truncatulinoides* (from V20–175) by addition of successive harmonics. The original shape is reproduced using the first 20 harmonics.

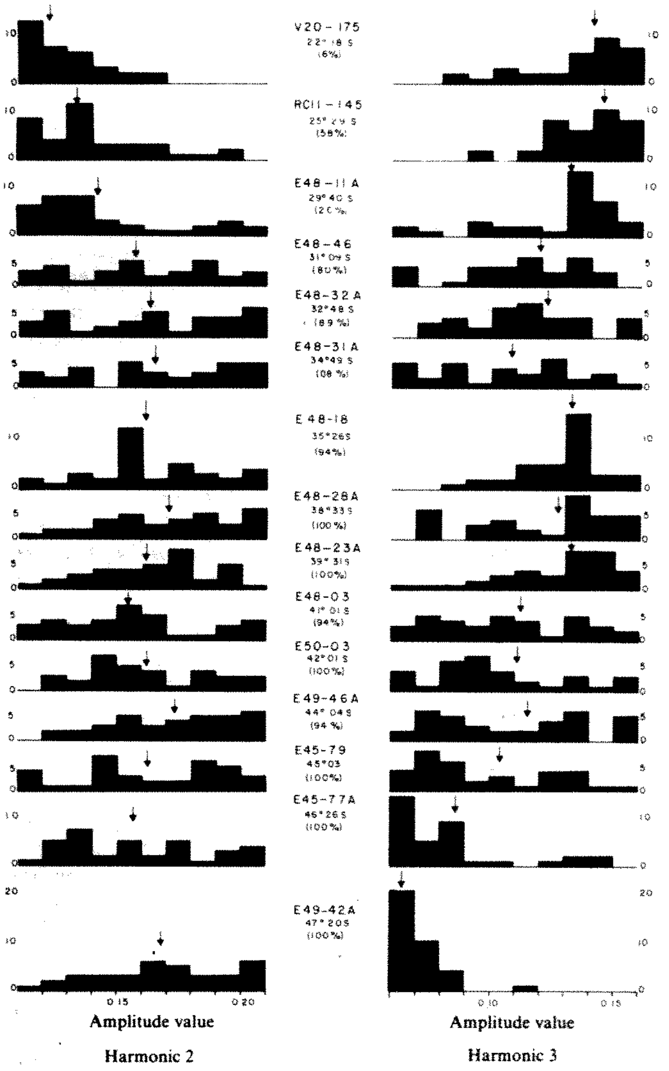


Fig. 3 Histogram plots of harmonic amplitude values for harmonic 2 (elongation) and harmonic 3 (triangularity/conical shape) in *G. truncatulinoides*. The histograms are ranked according to latitude and the per cent sinistrally coiled specimens in each sample is also shown. Arrows indicate the mean harmonic amplitude values in each sample.

variation (that is genetic and environmental influence) than ecophenotypic variation if we assume that coiling direction is genetically controlled. The factors governing coiling in foraminifera are not well understood[2]. However, the coiling of *G. truncatulinoides* in the southern Indian Ocean may reflect more of a clinal relationship rather than a strict environmental relationship, because this species remains dominantly sinistrally coiled across such an important oceanographical and biogeographical boundary as the subtropical convergence.

This study clearly illustrates that the conical shape of *G. truncatulinoides* varies as a function of water temperature and, as such, can be used as a palaeoceanographical index of water temperature in this region of the southern Indian Ocean. Current work uses the conicalness/ water temperature correlation to estimate water mass changes in the southern Indian Ocean sediments during the late Quaternary. The significance of the relationship between test elongation and coiling direction can only be resolved by studies of living plankton and laboratory-cultured specimens. We have demonstrated, however, that Fourier shape analysis is an accurate and quantitative method for rapid determination of the components of test shape in foraminifera. Because test shape may vary as a function of phylogenetic (evolutionary) affinities, environmental stress and ecophenotypic variation, we believe that Fourier shape analysis could be a powerful tool in morphometric studies of Cenozoic and late Mesozoic foraminifera.

We thank James P. Kennett for helpful criticism, Robert Ehrlich for helpful discussions, Cary Fico for his assistance with the Fourier software and Duane Eppler for use of his Drawshape program. This research was supported by NSF grant DPP77-21929 and a University of South Carolina Faculty and Productive Scholarship Grant. Belle W. Baruch Institute for Marine Biology and Coastal Research contribution no. 376.

Received 18 August; accepted 31 October 1980.

1. Scott, G. H. in *Foraminifera* Vol. 1 (eds Hedley, R. H. & Adams, C. G.) 55–151 (Academic, London, 1974).
2. Kennett, J. P. in *Foraminifera* Vol. 2 (eds Hedley, R. H. & Adams, C. G.) 111–170 (Academic, London, 1976).
3. Berger, W. H. *J. Paleont.* **43,** 1369–1383 (1969).
4. Parker, F. L. *Micropaleontology* **8,** 219–254 (1962).
5. Hecht, A. D. *J. Foram. Res.* **6,** 295–311 (1976).
6. Hecht, A. D., Be, A. W. H. & Lott, L. *Science* **194,** 422–424 (1976).
7. Malmgren, B. & Healy-Williams, N. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **25,** 235–240 (1978).
8. Be, A. W. H. *Micropaleontology* **19,** 150–192 (1973).
9. Malmgren, B. A. *Mar. Micropaleont.* **1,** 3–25 (1976).
10. Gradstein, F. M. *Mediterranean Pliocene Globorotalia: A Biometrical Approach*, 1–128 (Odijk, The Netherlands, 1974).
11. Younker, J. L. & Ehrlich, R. *Syst. Zool.* **26,** 336–341 (1977).
12. Kaesler, R. L. & Waters, J. A. *Bull. geol. Soc. Am.* **83,** 1169–1178 (1972).
13. Anstey, R. L. & Delmet, D. A. *Science* **177,** 1000–1002 (1972).
14. Anstey, R. L. & Delmet, D. A. *Bull. geol. Soc. Am.* **84,** 1753–1764 (1973).
15. Christopher, R. A. & Waters, J. A. *J. Paleont.* **48,** 697–709 (1974).
16. Kennett, J. P. *Science* **159,** 1461–1463 (1968).
17. Takayanagi, Y., Niitsuma, N. & Sakai, T. *Sci. Rep. Tohoku Univ. (Geol.)* **40,** 141–170 (1968).
18. Ehrlich, R. & Weinberg, B. *J. sed. Petrol.* **40,** 205–212 (1970).
19. Christopher, R. A. & Waters, J. A. *J. Paleont.* **48,** 697–709 (1974).
20. Gordon, A. L. in *Antarctic Oceanology II, Ant. Res. Ser.* Vol. 19 (ed. Hayes, D. E.) 3–9 (Am. Geophys. Union, Washington, 1972).
21. Williams, D. F. thesis, Univ. Rhode Island (1976).
22. Prell, W. L., Hutson, W., Williams, D. F. & Be, A. W. H. *Quat. Res.* (in the press).
23. Healy-Williams, N. (in preparation).
24. Benzecri, J. P. in *L'Analyse des Correspondances 619* (Dunod, Paris, 1973).

Parental investment in male and female offspring in polygynous mammals

T. H. Clutton-Brock, S. D. Albon & F. E. Guinness

King's College Research Centre, Cambridge CB2 1ST, UK*

Two evolutionary theories predict how mammalian parents would be expected to allocate the resources at their disposal to their male and female progeny. Where reproductive success varies more widely among males than females and variation in success among adults is influenced by parental investment[1], parents would maximize their reproductive success by allocating a greater proportion of their resources to individual sons than to individual daughters[2–5]. However, because the benefits of producing offspring of one sex are inversely related to the total investment allocated to them, parents should, on average, divide their total investment equally between their male and female progeny, rearing fewer of whichever sex is individually more expensive to produce[6–10]. Here we examine the extent to which parental investment in red deer (*Cervus elaphus*) and other polygynous mammals matches these predictions. We conclude that, in several mammals, mothers invest more heavily in individual sons than daughters but that, contrary to prediction, there is no indication that fewer male offspring are reared in these species.

In most polygynous mammals, lifetime reproductive success probably varies more widely among males than females and is affected by maternal investment[1]. For example, in the red deer of the Isle of Rhum (Scotland), where hinds live for 12–14 yr, mating in October, giving birth in June and weaning their calves

---

* Present address: University of Cambridge, Department of Zoology, Large Animal Research Group, 34A Storey's Way, Cambridge CB3 0DT, UK.

0028-0836/81/060487—03$01.00

Table 1 Birth sex ratios of calves born between 1970 and 1979 to different categories of red deer hinds

Age of mother (yr)	3	4	5–6	7–8	9–10	11–12	13+
% Males	52.6	59.7	56.0	58.3	53.8	58.8	69.0
n	38	67	109	84	65	57	29

Reproductive status in previous year	Failed to produce calf	Calf died within month of birth	Calf reared through first winter
% Males	59.3	60.3	56.4
n	59	68	218

Quality of home range	Superior	Inferior
% Males	56.0	57.1
n	168	98

For details of data collection and allocation see ref. 13.

approximately 6 months later[11–13], individual mothers rear from 0 to 12 calves surviving to 1 yr old during their lifetime[14]. In contrast, stags, which hold large harems of hinds during the mating season, father from 0 to >25 calves surviving to 1 yr during their lifetime[14,15]. In both sexes, reproductive success is probably affected by maternal investment during the growth period, but the relationship is particularly close among stags, where growth during the first year of life is related to adult body size and fighting ability[14].

In precisely these circumstances, mothers should maximize their reproductive success by investing more heavily in individual sons than daughters. Do they do so? One way they might achieve this would be to vary the sex ratio of their progeny according to their own body condition, producing relatively more male offspring when they were in good condition (and could afford to invest heavily) and relatively fewer males when in poor condition, perhaps by selectively aborting or resorbing male fetuses[2]. However, there is little evidence that factors which affect maternal condition in red deer have an important influence on birth sex ratios, which do not differ significantly between mothers of different ages, between those that have suffered the energetic strains of rearing a calf through the previous winter and those that have not done so, or between those occupying superior compared with inferior home ranges, although sample sizes are admittedly small (Table 1). Moreover, despite claims to the contrary[2], evidence for an association between maternal condition and birth sex ratio in other mammals is ambiguous[16,17]: some studies have found a tendency for females in good condition to produce relatively more males[18–23], while others have found the reverse[24–27].

An alternative possibility is that the sex ratio of progeny does not vary with maternal condition but that mothers transfer a greater proportion of their body reserves to sons than to daughters[3–5]. Several lines of evidence suggest that this is the case in red deer: compared with females, male calves have significantly longer gestation lengths (236.1 days versus 234.2 days, $P < 0.05$), heavier birth weights (6.72 kg versus 6.23 kg, $P < 0.002$) and, by analogy with other studies of dimorphic deer[28–31], probably gain weight faster. Observations show that male calves suck from their mothers more frequently (5.1 ($n = 17$) versus 2.6 ($n = 10$) daily suckling bouts: $P < 0.05$) and there is a trend for the duration of suckling bouts involving male calves to be longer than those involving females (Fig. 1). Evidence that the costs of rearing male calves exceed those of rearing females are provided by comparisons of the subsequent reproductive performance of mothers: between 1971 and 1980, hinds that had reared a male calf through the previous winter were less likely to produce a calf than those which had reared a female (Wilcoxon matched pairs signed ranks test across years, $T = 11$, P (one-tailed) $= 0.05$). In addition, hinds calved on average 11 days later following winters when they had supported a male calf compared with those when they had supported a female (Wilcoxon matched pairs signed ranks test across individual hinds, $z = 2.66$, $n = 32$, $P < 0.01$). As calving date is known to be related to body condition in red deer[32,33] and late calving increases the chances that calves will die during their first winter[34], this provides additional evidence of the increased costs of rearing male calves.

Similar data for other species are uncommon. However, detailed studies of maternal investment in elephant seals (*Mirounga angustirostris*) have found that male pups are weaned later than females[3] and, age for age, tend to be larger, while in a variety of other polygynous mammals, males commonly show longer gestation lengths, heavier birth weights, faster growth rates and greater food requirements than females[4,35–42]. There is also some evidence that the energetic costs of producing and rearing males may exceed those of producing and rearing females: in sheep, lambs born co-twin with males are lighter than those born co-twin with females[43] and, in Indian cattle, two studies have found a (non-significant) tendency for the next conception to be delayed longer after a bull calf has been produced[44,45]. However, a contrasting situation is found in colonies of macaques, where dominant females who have produced daughters are significantly less likely to produce another offspring in the next season than those that have reared sons[46]. One possible explanation is that, in this case, parental investment may have a greater effect on the breeding success of daughters and less on that of the sons because daughters generally inherit their mother's dominance rank while male body size may have little effect on competitive success[47].

In species where sons cost more to produce and rear than daughters, the sex ratio of progeny should be biased against males by the age at which parental investment is terminated[8]. However, sex ratios at weaning (as well as at birth) are often male-biased in polygynous species[47]. In particular, slightly male-biased weaning ratios occur in both species for which there is good evidence of increased investment in males before weaning (Table 2).

There are at least two ways in which these results could be reconciled with current theory: (1) Parental investment in daughters (but not sons) continues after weaning, balancing heavier pre-weaning investment in sons. This may be the case in a number of species where daughters adopt territories or home ranges overlapping those of their mothers while sons disperse[4,49]: if the presence of resident daughters reduces the amount of food available to the mother and her subsequent offspring, maternal tolerance of mature offspring should be regarded as a continuation of parental investment. However, this argument is unlikely to explain the existence of male-biased weaning ratios in either red deer or elephant seals. In the latter, both sexes of offspring are abandoned by their mothers on the breeding beaches and there is no evidence of social bonds between mothers and their weaned offspring[3,48]. In the former, toleration of mature daughters is unlikely to represent an important form of investment since unrelated hinds occupy overlapping home ranges, so that if a mother were to exclude her daughters from her range, this would have little effect on the resources available to her. (2) Alternatively, the primary sex

Table 2 Sex ratios at birth and after weaning in two polygynous mammals

	Sex ratio at:	
	Birth	12 months (~6 months post-weaning)
Red deer[14] (*Cervus elaphus*)		
% males	57.0	53.3
n	423	296
	Birth	1 month (shortly after weaning)
Northern elephant seal[48] (*Mirounga angustirostris*)		
% males	51.6	51.3
n	2,396	2,052

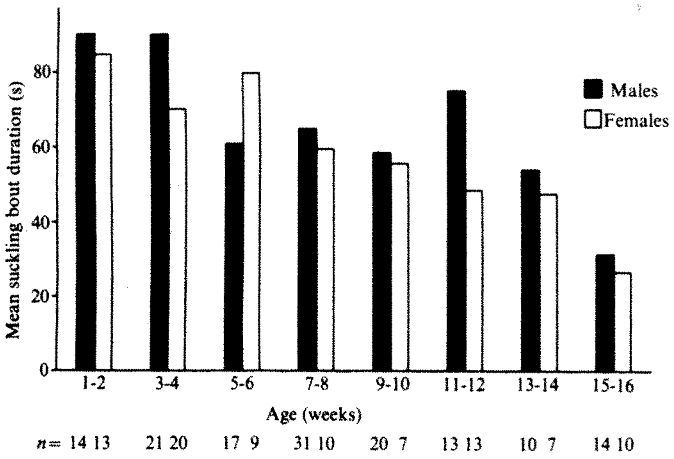


Fig. 1 Duration of suckling bout(s) to male and female calves between July and October.

ratio may be fixed at unity. In this case, selection for increased investment in individual sons could prefer parents that allocated a greater proportion of their total investment to male progeny[5]. However, evidence that sex ratios can be modified argues against this theory[47].

Thus, the available data suggest that polygynous mammals commonly invest more in individual sons than daughters and that this tendency may be associated with a greater total investment in male than female progeny. Why total investment is not equalized between males and females remains an anomaly.

We thank Rosemary Cockerill, Professor J. Maynard Smith, Professor R. V. Short and Dr Paul Harvey for advice and criticism and the NERC, SRC and Leverhulme Trust for support.

Received 23 May; accepted 20 November 1980.

1. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B. G.) 136-179 (Aldine, Chicago, 1972).
2. Trivers, R. L. & Willard, D. E. *Science* **179,** 90–92 (1973).
3. Reiter, J., Stinson, N. L. & LeBoeuf, B. J. *Behav. Ecol. Sociobiol.* **3,** 337–367 (1978).
4. Dittus, W. P. J. *Behaviour* **69,** 265–301 (1979).
5. Maynard Smith, J. *Behav. Ecol. Sociobiol.* **7,** 247–257 (1980)
6. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
7. MacArthur, R. H. in *Theoretical and Mathematical Biology* (eds Waterman, T. & Morowitz, H.) 388–397 (Blaidsdell, New York, 1965).
8. Leigh, E. G. Jr. *Am. Nat.* **104,** 205–210 (1970).
9. Eshel, I. *Heredity* **34,** 351–361 (1975).
10. Charnov, E. L. *Am. Nat.* **113,** 465–480 (1979).
11. Lowe, V. P. W. *J. Anim. Ecol.* **38,** 425–457 (1969).
12. Clutton-Brock, T. H. & Guinness, F. E. *Behaviour* **55,** 287–300 (1975).
13. Guinness, F. E., Albon, S. D. & Clutton-Brock, T. H. *J. Reprod. Fert.* **54,** 325–334 (1978).
14. Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. *Red Deer: The Ecology of Two Sexes* (Chicago University Press, in the press).
15. Clutton-Brock, T. H., Albon, S. D., Gibson, R. M. & Guinness, F. E. *Anim. Behav.* **27,** 211–225 (1979).
16. Myers, J. H. *Am. Nat.* **112,** 389–399 (1978).
17. Williams, G. C. *Proc. R. Soc.* **B205,** 567–580 (1979).
18. Coulson, J. C. & Hickling, G. *Nature* **190,** 281–282 (1961).
19. Stirling, I. *J. Mammal.* **52,** 842–844 (1971).
20. Davies, G. *Nature* **127,** 94–95 (1931).
21. Robinette, W. L., Gashwiler, J. S., Jones, D. A. & Crane, H. S. *J. Wildl. Mgmt* **19,** 115–136 (1955).
22. Rivers, J. P. W. & Crawford, M. A. *Nature* **252,** 297–298 (1974).
23. Teitelbaum, M. S. *J. biosoc. Sci.* Suppl. 2, 61–71 (1970).
24. Miller, W. C. *Proc. R. phys. Soc.* **22,** 99–101 (1932).
25. Verme, L. J. *J. Wildl. Mgmt* **33,** 881–887 (1969).
26. Mech, L. D. *J. Wildl. Mgmt* **39,** 737–740 (1975).
27. Andersen, F. S. *Oikos* **12,** 1–16 (1961).
28. Blaxter, K. L., Kay, R. N. B., Sharman, G. A. M., Cunningham, J. M. M. & Hamilton, W. J. *Farming the Red Deer* (HMSO, London 1974).
29. McEwan, E. A. & Whitehead, P. E. *Can. J. Zool.* **49,** 443–447 (1971).
30. Robbins, C. T. & Moen, A. N. M. *J. Wildl. Mgmt* **39,** 355–360 (1975).
31. Krebs, C. J. & Cowan, I. Mc.T. *Can. J. Zool.* **40,** 863–869 (1962).
32. Mitchell, B. & Lincoln, G. A. *J. Zool.* **171,** 141–152 (1973).
33. Guinness, F. E., Gibson, R. M. & Clutton-Brock, T. H. *J. Zool.* **185,** 105–114 (1978).
34. Guinness, F. E., Clutton-Brock T. H. & Albon, S. D. *J. Anim. Ecol.* **47,** 817–832 (1978).
35. Benedict, F. G. *Vital Energetics: A Study of Comparative Basal Metabolism* (Carnegie Institute, Washington, 1938).
36. Defries, J. C., Touchberry, R. W. & Hayes, R. L. *J. Dairy Sci.* **42,** 598–606 (1959).
37. Short, R. V. *J. Reprod. Fert.* **1,** 61–70 (1960).
38. Willis, M. B. & Wilson, A. *Anim. Prod.* **18,** 231–236 (1974).
39. Glucksman, A. *Biol. Rev.* **49,** 423–475 (1974).
40. McEwan, E. H. *Can. J. Zool.* **46,** 1023–1029 (1968).
41. Mitchell, B., Staines, B. V. & Welch, D. *Ecology of Red Deer* (Institute of Terrestrial Ecology, Cambridge, 1977).
42. Morrison, F. B. *Feeds and Feeding,* 21st edn (Morrison, New York, 1948).
43. Burfening, P. J. *Anim. Prod.* **15,** 61–66 (1972).
44. Dhillon, J. S., Acharya, R. M., Tiwana, M. S. & Aggarwal, S. C. *Anim. Prod.* **12,** 81–87 (1970).
45. Singh, O. N., Singh, R. N. & Srivastava, R. R. P. *Indian J. vet. Sci.* **35,** 245–248 (1965).
46. Simpson, M. J. A., Simpson, A. E., Hooley, J. & Zunz, M. *Nature* (in the press).
47. Clutton-Brock, T. H. & Albon, S. D. in *Current Problems in Sociobiology* (ed. King's College Sociobiology Group) (Cambridge University Press, in the press).
48. LeBoeuf, J. & Briggs, K. T. *Mammalia* **41,** 167–195 (1977).
49. Clark, A. B. *Science* **201,** 163–165 (1978).

Photon-like signals following weak rhodopsin bleaches

J. F. Ashmore* & G. Falk

Department of Biophysics, University College London,
Gower Street, London WC1E 6BT, UK

The rod visual system in its dark-adapted state behaves as a near-ideal light detector[1]. Psychophysical studies on the reliability of light detection in man[2,3], analysis of the dark noise in rod bipolar cells[4] and observation of photon-like events in rods[5] in the dark suggest that the visual pigment, rhodopsin, is very stable against spontaneous isomerization. When a light which bleaches a small fraction of the rhodopsin is extinguished, the visual threshold may be increased by several orders of magnitude. The eye is then 'light-adapted' and the process (or processes) by which the sensitivity returns constitutes dark adaptation. We report here that bleaching of a small fraction of the rhodopsin produces a prolonged increase in the noise observed in the dark in rod bipolar cells of the dogfish retina. The associated noise events are similar to those produced by the absorption of light quanta and presumably have their origin in the rods which transmit their signals to the bipolar cells. This increased noise after bleaching would decrease the reliability of detection of light quanta and contribute to the elevation of visual threshold.

Responses to light were recorded intracellularly in rod bipolar cells in the initially dark-adapted, hemisected eyecup of the dogfish, *Scyliorhinus canicula*. The bipolar cells studied here were those which in the dark-adapted state depolarized in response to light delivered as full-field illumination. The methods of recording, identification of cells and estimation of irradiance and bleaching have been described previously[6,7]. Because of the high voltage gain in signal transmission from rods to bipolar cells[7], we could study responses to light and changes in voltage noise during and following light which bleached fractions of between 10^{-9} and 10^{-3} of the rhodopsin.

Figure 1*a* shows the noise increase during dim light, bleaching, on average, about one rhodopsin molecule per rod in 50 s. The noise at higher gain and faster sweep during the period indicated by arrows is shown as the middle record of Fig. 2*a*. Power-spectral analysis of this noise suggests that most of the noise arises from random absorption of light quanta by the rods within the receptive field of the bipolar cell ('photon noise')[4]. When the dim light was switched off, there was no detectable change in the dark noise.

Figure 1*b* shows the response of the same cell to light 4,000 times brighter, during which time the dark noise was suppressed

---

* Present address: Ethology and Neurophysiology Group, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK.

0028-0836/81/060489—03$01.00

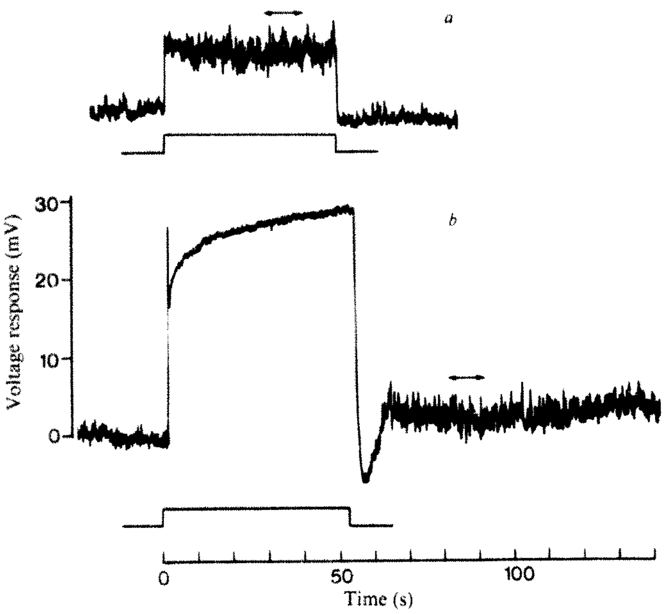


Fig. 1 Voltage responses of a rod bipolar cell to steps of light, accompanied by an increase of noise during very dim illumination (*a*) and a prolonged increase of noise following a bleach (*b*); both records are from the same cell. The duration of the light stimulus is indicated under each record, which in all experiments was delivered as full-field illumination. *a*, The total bleach during the step was 1.1 Rh**, at a rate of 0.022 Rh** s^{-1}, where Rh** signifies a rhodopsin molecule bleached per rod[7]. The voltage noise variance in the dark (corrected for noise due to the electrode determined with the electrode in the cell in a bright light) was 0.20 mV^2; the noise variance during illumination was 0.95 mV^2. *b*, The response to a step of light with a total bleach of 4,300 Rh** at a rate of 79 Rh** s^{-1}. Assuming 3×10^8 rhodopsin molecules per rod (ref. 4), this corresponds to a 0.0014% bleach. Note that the noise increased with a delay following the light step. The noise variance following the bleach (during the period indicated by arrows) was 1.0 mV^2. Potential of the cell in the dark, −56 mV; maximum response to a bright flash, 30 mV; flash sensitivity 1,860 mV per Rh** (ref. 7); step sensitivity, 645 mV s per Rh** (ref. 7). Temperature, 17.5 °C. An increase in noise was observed in 7 out of 10 cells following light steps producing bleaches of 10^{-6}–10^{-5} of the rhodopsin. Low sensitivity and noise in the recording system may have obscured the effect in some cells.

and then increased with a delay of about 10 s when the light was switched off. Several features of Fig. 1*b* require comment. The non-steady depolarization during light may be due to a bipolar cell membrane non-linearity, while the slow, upward creep of the depolarization during steady light may indicate a slow inactivation of potassium conductance. Although this membrane non-linearity may contribute to a decrease in noise during light which produced response saturation, the decrease in noise can be shown to be due mainly to the non-linear relation between light intensity and depolarization. Only in the linear region of response does photon noise variance increase in proportion to the mean depolarization. As saturation of the response is approached, the noise variance declines to zero. In Fig. 1*b*, the light bleached 4,300 rhodopsin molecules per rod or about 0.0014%. When it was turned off, there was a transient hyperpolarization followed by a prolonged after-depolarization, accompanied by an increase in noise variance. The after-depolarization reached its peak value of 5 mV and the noise variance increased to six times its dark value 10 s after the light was turned off. The noise variance returned to its initial dark value in about 3 min. The post-bleach noise during the interval indicated by arrows in Fig. 1*b* is shown at an expanded sweep and amplification in Fig. 2*a* (bottom record). Clearly, the character of the post-bleach noise is similar to that during dim light and this similarity to photon noise was confirmed by their power spectra.

The effect of light bleaching 0.17% rhodopsin is illustrated in Fig. 3, recorded in another cell. As in the cell of Fig. 1, the noise increased after a delay and the increased noise was associated with a depolarization lasting about 10 min. The noise recorded about 4.5 min after the light was switched off is shown at an expanded sweep and higher gain as the bottom record of Fig. 2*b*. The post-bleach noise had characteristics similar to the noise recorded during dim light in the same cell (the middle record of Fig. 2*b*). The power spectra of the noise recorded at different times during the period following the bleaching were all indistinguishable (except for a power-scale factor) and were the same as during dim light. Because the events giving rise to the increased noise are like photon events, we presume that they arise in the rods. Recently, Lamb has observed increased fluctuations in single rods, resembling light quantal events, which are induced by bleaching[8].

We can estimate the equivalent rate of photon-like events in the rod in the following way. For a linear shot process model for

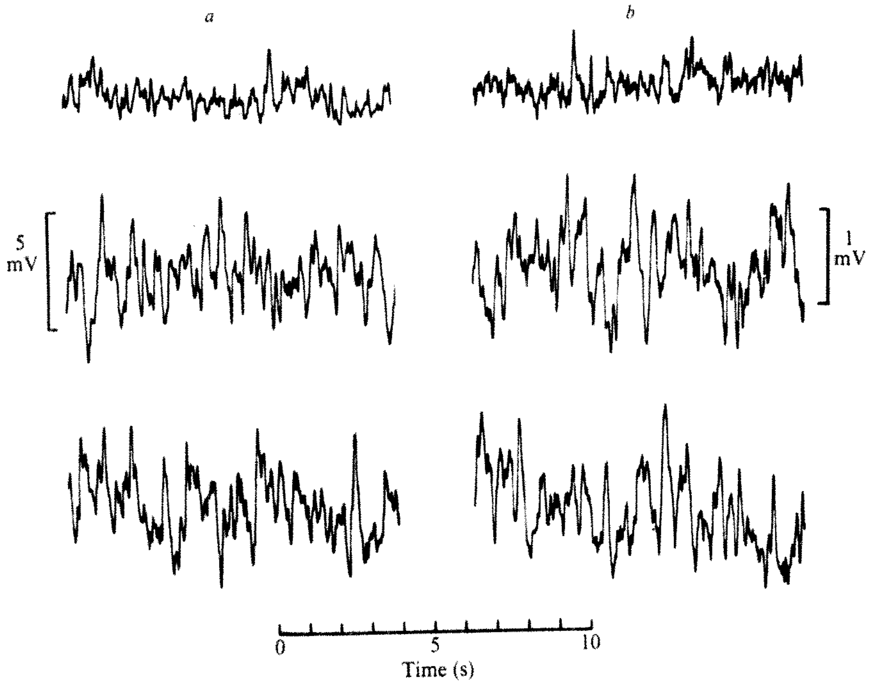


Fig. 2 Voltage noise in bipolar cells in darknesss (upper row), during dim illumination (middle row) and following bleaches (bottom row), at higher gain and sweep speeds. *a*, Records from the cell illustrated in Fig. 1. The middle and lower records are from the intervals indicated by the arrows in Fig. 1*a* and Fig. 1*b*, respectively. *b*, Records from the cell illustrated in Fig. 3. The middle record shows the noise during dim illumination at an intensity of 0.026 Rh** s^{-1} (~1 molecules bleached per 40 rods per s), which depolarized the cell by 2.7 mV. The bottom record shows the noise following a 0.17% bleach, corresponding to the interval arrowed in Fig. 3.

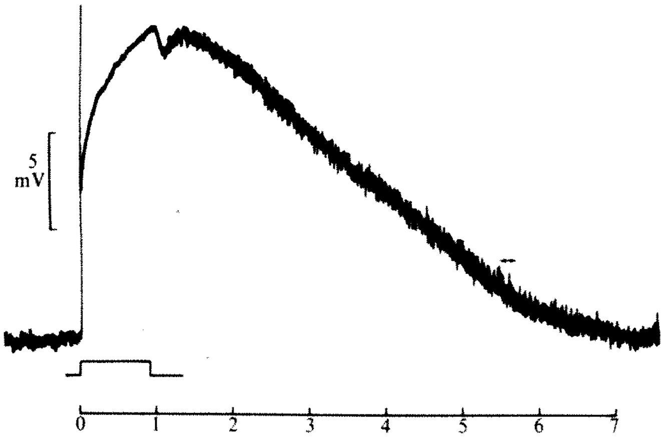


Fig. 3 Prolonged increase of noise in a rod bipolar cell following a pigment bleach of 0.17%. The rate of bleaching was 9.4×10^3 Rh** s^{-1} so that the total bleach was 5.2×10^5 Rh**. The noise variance in the dark (corrected for electrode noise) was 0.01 mV^2. Approximately 4.5 min after the bleaching light was turned off, the noise variance was 0.074 mV^2. The cell was held for 15 min following the bleach during which time the potential and noise returned to their pre-bleach values. Potential of the cell in the dark, −29 mV; maximum response to light, 17 mV; flash sensitivity, 320 mV per Rh**; step sensitivity, 120 mV s per Rh**. Temperature, 17 °C.

events seen in the bipolar cell, the equivalent rate per rod would be

$$\lambda = V/(aN\tau)$$

where V is the mean depolarization, a is the amplitude of the shot event in the bipolar cell, N is the number of rods in its pool and τ is the integration time. (If, like the effect of light, the relation between mean depolarization and the rate of events is a rectangular hyperbola, λ would need to be corrected by a factor, $\{1-(V/V_{\max})\}^{-1}$, where $V_{\max}$ is the maximum response.) The product, $aN\tau$, is the step sensitivity and was determined in the well dark-adapted state as the slope of the relation between steady-state depolarization and light intensity in the linear range of response[7]. For the cell of Fig. 1, the peak rate of events, following a bleach of 0.0014% giving a 5-mV depolarization, was 0.01 per s per rod. This was 10.3 times the initial dark-adapted rate of equivalent events, determined in the way described by Ashmore and Falk[4]. Another bleaching exposure in the same experiment, bleaching 1/15th of that illustrated in Fig. 1*b*, resulted in an increased post-bleach noise variance of 70% greater than that of the initial well dark-adapted state. For the bleach of 0.17% illustrated in Fig. 3, the peak post-bleach rate (corrected for non-linear summation) amounted to 1.8 events per s per rod, which was 1,500 times the initial dark-adapted rate. These results suggest that the peak extra rate of noise events induced by bleaching is proportional to the amount of rhodopsin bleached, at least for small bleaches. Quantitatively,

$$\frac{\lambda_B}{\lambda_D} = \alpha R$$

where λ_B is the peak rate of additional events induced by bleaching, λ_D is the dark-adapted rate of equivalent events, R is the number of rhodopsin molecules bleached per rod and α the constant of proportionality amounting to $2\text{–}3\times 10^{-3}$ per Rh**, where Rh** denotes one rhodopsin molecule bleached per rod.

These relatively weak bleaches give rise to an increased rate of spontaneous events lasting many minutes and the question arises of how the total number of extra events is related to the number of rhodopsin molecules bleached. Integration of the increased rate over time indicates that bleaching of a rhodopsin molecule has only a small probability, less than 0.001, of inducing an event. It is unlikely that this low probability is the result of regeneration of rhodopsin which in the elasmobranch eye[9] is very much slower than the rate of recovery of the noise to its dark-adapted value.

The present results confirm Rushton's[10] and Barlow's[11] conjecture that bleaching of rhodopsin in rods induces signals which constitute a background of 'dark light' or noise which would raise the threshold for detection of a light stimulus. Rushton[10] found that the equivalent background light increased exponentially with the fraction of rhodopsin in the bleached state. Extrapolated to weak bleaches, the exponential relationship becomes one of linearity. However, note that the value of α reported here for weak bleaches is several orders of magnitude greater than that inferred from the psychophysical measurements[10,11] with strong bleaches. Likewise, the rise in visual threshold in man following weak bleaches[12,13] is also greater than that predicted from near-total bleaching.

An increase in the rate of occurrence of discrete events following bleaching may be general among photoreceptors for it has been reported to occur in the photoreceptors of invertebrates[14,15]. We cannot decide yet whether the increased rate of spontaneous events in rods following absorption of light quanta is due to an increased rate of thermal isomerization of rhodopsin or to a link with some reaction involved in phototransduction.

The work was supported by an MRC project grant.

Received 29 August; accepted 19 November 1980.

1. Hecht, S., Shlaer, S. & Pirenne, M. H. *J. gen. Physiol.* **25**, 819–840 (1942).
2. Denton, E. J. & Pirenne, M. H. *J. Physiol., Lond.* **123**, 417–442 (1954).
3. Barlow, H. B. *J. opt. Soc. Am.* **46**, 634–639 (1956).
4. Ashmore, J. F. & Falk, G. *Nature* **270**, 69–71 (1977).
5. Yau, K.-W., Lamb, T. D., Matthews, G. & Baylor, D. A. *Vision Res.* **19**, 387–390 (1979).
6. Ashmore, J. F. & Falk, G. *Nature* **263**, 248–249 (1976).
7. Ashmore, J. F. & Falk, G. *J. Physiol., Lond.* **300**, 115–150 (1980).
8. Lamb, T. D. *Nature* **287**, 349–351 (1980).
9. Dowling, J. E. & Ripps, H. *J. gen. Physiol.* **56**, 491–520 (1970).
10. Rushton, W. A. H. *J. opt. Soc. Am.* **54**, 104–109 (1963).
11. Barlow, H. B. *Vision Res.* **4**, 47–58 (1964).
12. Rushton, W. A. H. & Powell, D. S. *Vision Res.* **12**, 1073–1081 (1972).
13. Pugh, E. N. *J. Physiol., Lond.* **248**, 413–432 (1975).
14. Fein, A. & Hanani, M. *Brain Res.* **156**, 157–161 (1978).
15. Tsukahara, Y. & Horridge, G. A. *J. exp. Biol.* **68**, 137–149 (1977).

Isolation of the phagocytosis-inducing IgG-binding antigen on senescent somatic cells

Marguerite M. B. Kay

Laboratory of Molecular and Clinical Immunology, Geriatric Research, Education and Clinical Center, Veterans Association Wadsworth Medical Center, and the Department of Medicine, University of California, Los Angeles, California 90073

To remove senescent red blood cells (RBCs) from the circulation, macrophages must distinguish them from mature RBCs. That is achieved by a specific recognition system[1,2]. An antigen that develops on the surface of a senescing RBC is recognized and bound by the Fab region[1] of an IgG autoantibody in the serum[2]. Subsequently the Fc region of the autoantibody is recognized and bound by a macrophage[3], which proceeds to phagocytose the RBC. The antigenic molecule can be extracted from senescent but not young RBCs with Triton X-100 (ref. 4), although 10–30% as much antigen can be extracted from middle-aged as from senescent RBCs[4]. I have now used IgG autoantibodies eluted from senescent RBCs to isolate and purify the IgG-binding antigen on senescent RBCs, and to detect the antigen on other somatic cells. The antigen is a ≃62,000-M_r protein which is present on stored platelets, lymphocytes and neutrophils, and on cultured human adult liver and embryonic kidney cells, as well as senescent RBCs.

0028-0836/81/060491—04$01.00

IgG was eluted from human senescent RBCs[5], by the method of Kochwa and Rosenfield[6] (for details see Fig. 1). The IgG-containing eluates of senescent RBCs from 50 l of blood were pooled, and IgG isolated from the eluates by affinity chromatography with Protein A–Sepharose 4B (Pharmacia)[1].

Glycine–HCl eluates from the affinity column contained IgG and an additional peptide that was not part of the IgG molecule, as determined by polyacrylamide gel electrophoresis on gradient gels (Fig. 1). This peptide has a M_r of ≈62,000 and migrates between IgG heavy chain and its dimer on polyacrylamide gels and is not present in IgG isolated from serum with Protein A columns (Fig. 1). This peptide stains with dansyl hydrazine and thus is probably a sialoglycopeptide (not shown). Elution of the ≈62,000-M_r peptide from Protein A columns along with the IgG obtained from senescent cells suggests that it may be the senescent cell antigen for IgG autoantibodies.

As the peptide that co-purifies with the IgG eluted from senescent cells seems to be a sialoglycopeptide[1,2,4], I attempted to isolate the senescent cell antigen for IgG autoantibodies from preparations of RBC sialoglycoproteins. IgG eluted from senescent cells and purified by affinity chromatography with

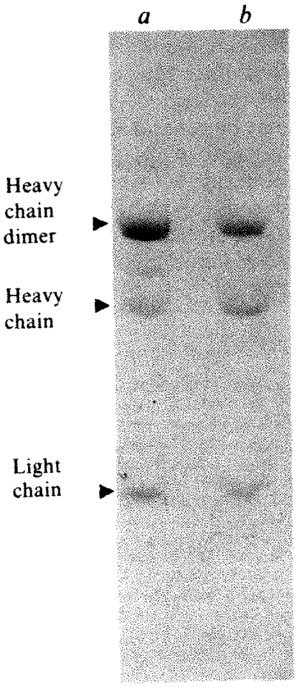


Fig. 1 Polyacrylamide gel electrophoresis of IgG eluted from senescent RBCs and purified with Protein A affinity columns. *a*, IgG eluted from senescent RBCs and purified with a Protein A–Sepharose column; *b*, IgG isolated from normal human serum by ion-exchange chromatography followed by Protein A–Sepharose chromatography. Blood was obtained from 250 healthy individuals between the ages of 21 and 35 yr, and was processed the same day that it was obtained. The blood was centrifuged for 15 min at 2,000*g* and the serum and white cells removed. RBCs were separated into populations of different ages by centrifugation[4,5,8], and young RBCs were removed. The remaining RBCs were washed three times with 50–100 vol of Dulbecco's phosphate-buffered saline (PBS), *p*H 7.4. IgG was eluted as described previously[1,6]. Briefly, RBC membranes (ghosts) were prepared by digitonin lysis, washed three times with PBS, and IgG eluted with 0.1 M glycine–HCl buffer, *p*H 2.3. Eluates were neutralized with 1 M NaOH and concentrated to 0.25–1.00 ml using an Amicon Diaflow with a PM 10 filter. Solid material was removed by centrifugation at 27,000*g*. Samples were stored at −80 °C until use. They were then thawed, pooled and incubated with Protein A–Sepharose 4B for 30–60 min at 37 °C and 24–48 h at 4 °C. The Protein A–Sepharose was poured into a 12-ml column prepared from a syringe and washed with 100 vol of PBS, then 100 vol of 10 × PBS followed by 100 vol of PBS. IgG was eluted with 2 vol of glycine–HCl buffer, *p*H 2.3, neutralized with 1 M NaOH and dialysed against distilled water (*a*). The amount of IgG obtained was 7.42 mg. Polyacrylamide gel electrophoresis was performed with 2.7-mm thick gradient slab gels containing 4–30% acrylamide. The buffer system of Laemmli was used[9]. Gels were stained for glycoproteins with dansyl hydrazine following the method of Eckhardt *et al.*[10] with the modifications of Lutz *et al.*[11]. Proteins were stained with Coomassie blue (shown here). Serum IgG was isolated by ion-exchange chromatrography on DEAE-cellulose (DE 52, Whatman)[1,2] followed by incubation with Protein A–Sepharose as described above (*b*).

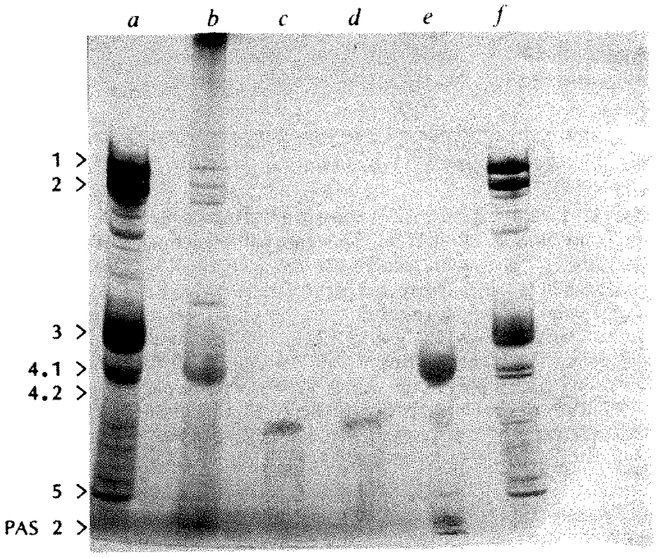


Fig. 2 Isolation of the antigen for IgG autoantibodies from human RBC sialoglycoproteins with affinity columns prepared from the IgG eluted from senescent RBCs. *a*, RBC ghost; *b*, Triton X-100 extract of ghosts; *c*, 62,000-M_r peptide isolated with senescent cell IgG affinity column from Triton X-100 extract; *d*, 62,000-M_r peptide isolated with the senescent cell IgG affinity column from lithium diiodosalicylate (LIS) extract of ghosts; *e*, LIS extract of ghosts; *f*, RBC ghost. Blood was obtained from healthy donors 21–30 yr old. Donor cells were typed for 27 known RBC blood group antigens, including ABO, Rh, MN, Ss̄, Duffy, Kell and Kidd. Donors did not have detectable autoantibodies to RBCs or antibodies to any known blood group antigen, as determined by the Immunohematology Laboratory, American Red Cross, nor did they have antibodies to nuclei, DNA, mitochondria, parietal cells or smooth muscle[12]. All blood was processed the same day that it was obtained. RBCs from 500 ml of human blood were depleted of white cells and young RBCs by centrifugation on self-forming Percoll gradients[4,8]. RBCs were washed three times in 50–100 vol of PBS with 75 μg ml^{-1} phenylmethylsulphonylfluoride (PMSF) at 4 °C. Right-side out RBC membranes were prepared by hypotonic lysis in 5 mM sodium phosphate buffer, *p*H 7.5, containing 50 μg ml^{-1} PMSF[11]. After five washes in 100 vol of the same buffer (*a* and *f*), ghosts were resuspended in a solution of 0.3 M LIS[7] or 1.0% Triton X-100 and partitioned in a phenol–water mixture according to the procedure of Marchesi and Andrews[7]. The water phase was dialysed extensively against distilled water at 4 °C and then lyophilized (*b*). The lyophilized material extracted with LIS was resuspended in 5 mM phosphate buffer and extracted with chloroform–methanol (2 : 1)[13]. The aqueous phase was dialysed against distilled water (*e*) and processed with the senescent cell IgG affinity column. Triton X-100 extracts were processed with the affinity column without chloroform–methanol treatment. The affinity column was washed with 100 vol of the following: PBS with 50 μg ml^{-1} PMSF, 10 × PBS, 5 mM phosphate buffer, and then PBS with 50 μg ml^{-1} PMSF. Material bound by the column was eluted with 15 vol of glycine–HCl, *p*H 2.3, neutralized, dialysed against distilled water at 4 °C and lyophilized (*c* and *d*). Electrophoresis was performed using gradient gels containing from 2 to 16% acrylamide and the buffer system of Fairbanks *et al.*[14], except that 0.2% SDS was used. Sample buffer: 10 mM Tris-HCl, *p*H 8.0, containing 1 mM Na EDTA, 40 mM dithiothreitol, 5% glycerol, 0.2% SDS and bromophenol blue. All of the proteins extracted with LIS stained with dansyl hydrazine. The photograph shows the gel stained with Coomassie blue. PAS, periodic acid-Schiff: designation for RBC sialoglycoproteins.

Protein A was conjugated to cyanogen bromide-activated Sepharose 4B (Pharmacia). Columns were prepared and washed with 15 vol of glycine–HCl buffer, *p*H 2.3, to break the bond between the insolubilized IgG and its senescent cell antigen. After neutralization with NaOH, preparations of pure RBC membrane sialoglycoproteins, obtained by extraction with lithium diiodosalicylate and partition in a phenol–water mixture according to the procedure of Marchesi and Andrews[7] (see Fig. 2 for details), were processed with the senescent cell IgG column. After extensive washing, the material on the column was eluted with glycine–HCl buffer, *p*H 2.3, neutralized, dialysed against deionized, distilled water at 4 °C, lyophilized and analysed by gel electrophoresis. Both dansyl hydrazine and Coomassie blue stains of the gels revealed a single band migrating at a M_r of ≈62,000 which appears below band 4.2 in Fig. 2. When this experiment was repeated using 1% Triton X-100 extracts of RBC membranes[8], the same results were obtained. A single peptide of $M_r \approx 62,000$ was observed on gels of the material eluted from the senescent cell IgG affinity column (Fig. 2). These experiments indicate that the ≈62,000-M_r peptide carries the antigenic determinants recognized by IgG obtained from freshly isolated senescent cells.

To confirm this finding, an erythrophagocytosis inhibition assay was performed[1,2]. IgG was absorbed by incubation at 37 °C for 30 min with either the 62,000-M_r peptide eluted from senescent cell IgG affinity columns or with the sialoglycoprotein mixture that had not bound to these affinity columns. The absorbed IgG was then incubated with stored RBCs[4] for 30 min at 37 °C because stored RBCs have the exposed receptor[1,2,4]. The RBCs were washed three times and incubated with autologous macrophage cultures[1,2,4]. RBCs incubated with the IgG absorbed with excess 62,000-M_r peptide were not phagocytosed (0% phagocytosis), indicating that no free antigen-binding IgG remained in solution. In contrast, 52±3% (±s.d.) of the RBCs incubated with IgG absorbed with the eluate that was not bound by senescent cell affinity columns were phagocytosed, indicating the presence of free antigen-binding IgG. The same amount of phagocytosis was observed when stored RBCs were incubated with IgG that was not absorbed (55±6%). Thus, the 62,000-M_r peptide, but not the remaining sialoglycoprotein mixture from which it was isolated, abolishes the phagocytosis-inducing ability of IgG. This indicates that the 62,000-M_r peptide is the antigen which emerges on the surface membrane of cells as they senesce.

Table 1 Inhibition of the phagocytosis-inducing ability of IgG autoantibody eluted from senescent RBC by previous absorption with *a*, stored platelets, lymphocytes or neutrophils, and *b*, human adult liver cells or human embryonic kidney cells

	Cell type used for absorption	% Phagocytosis (mean±s.d.)
a	None	37±4
	Platelets	2±2
	Lymphocytes	0
	Neutrophils	0
b	None	38±6
	Liver	0
	Kidney	7±7

In *a*, macrophages, lymphocytes, neutrophils and platelets were isolated by centrifugation on Percoll gradients[4]. Macrophages were cultured, and lymphocytes and platelets were stored in Alpha minimal essential medium with 10% fetal calf serum (FCS) at 4 °C (refs 2, 4). Citrate phosphate–dextrose solution was added to the platelet suspension to prevent activation. The specific IgG autoantibody eluted from senescent cells (6 μg) was added to 1-ml suspensions of ≈2×10^7 lymphocytes or 1×10^8 platelets and incubated for 30 min at 37 °C and 30 min at 24 °C (refs 1, 2, 4). Positive controls consisted of autoantibody incubated with medium only. Cells were removed by centrifugation and the supernatant containing the remaining antibody was added to 5×10^8 stored RBCs. RBCs were incubated for 15 min at 37 °C and 1.5 h at 24 °C. Aliquots of each sample (0.1 ml) containing 5×10^7 RBCs were added to macrophage cultures. Phagocytosis was determined as described previously[1,2,4,8]. Data are presented as the mean±1 s.d. The amount of IgG antibody used for absorption and the number of cells used were based on preliminary dose–response experiments. The phagocytosis-inducing ability of the senescent cell IgG could not be absorbed with freshly isolated young RBCs (43±9% phagocytosis) but was abolished by absorption with stored RBCs (0% phagocytosis). Data are presented as the mean±s.d. In *b*, human adult liver and human embryonic kidney (Chang and Flow 4000 respectively, Flow) cells were stored in liquid nitrogen, thawed and grown in tissue culture for 7 days before use. The media were changed 2 days after initiation of culture. The less viable non-adherent cells were collected by centrifugation and resuspended in 1 ml of Alpha minimum essential medium containing 10% FCS[4]. IgG eluted from senescent cells (3 μg) was incubated with 2.1×10^6 liver cells (57% did not exclude Trypan blue) or 5×10^6 kidney cells (82% did not exclude Trypan blue) for 20 min at 37 °C and 2 h at 24 °C. Cells were removed by centrifugation, and the supernatant incubated with stored RBCs. The phagocytosis inhibition assay was performed as described in the text.

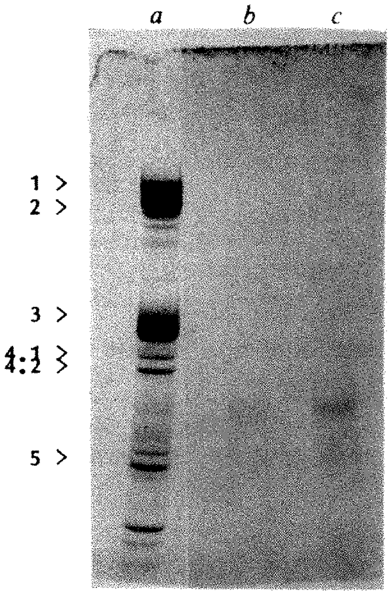


Fig. 3 Isolation of the antigen for IgG autoantibodies from human lymphocyte sialoglycoproteins with affinity columns prepared from the IgG eluted from senescent RBCs. *a*, RBC ghost; *b*, 62,000-M_r peptide isolated with the senescent cell IgG affinity column from LIS extract of RBC ghost; *c*, 62,000-M_r peptide isolated with the senescent RBC IgG column from LIS extract of lymphocyte membranes. Lymphocytes (8×10^8) were isolated from 500 ml of blood by centrifugation on Percoll gradients[4]. After three washes, membranes were prepared by hypotonic lysis in 5 mM NaPO$_4$ and 1 mM EDTA, *p*H 7.6 (ref. 15). Membrane suspensions were centrifuged at 248*g* to remove large particles. The supernatant containing the membranes was centrifuged at 27,000*g*. Membranes were washed three times in 100 vol of the same buffer, extracted with LIS and partitioned in a phenol–water mixture as described for RBCs in Fig. 2 legend. The aqueous phase was dialysed against distilled water and processed with the senescent RBC IgG affinity column.

Other somatic cells were then examined for the presence of the antigen which appears on senescent RBCs. Macrophages, lymphocytes, platelets and neutrophils were isolated by centrifugation on Percoll density gradients[4,9]. Cells were stored at 4 °C for 48 h. They were then washed and incubated with the autoantibody eluted from senescent cells (Table 1). Cells were removed by centrifugation and the supernatants containing any remaining IgG were incubated with stored RBCs. Stored RBCs were washed and incubated with macrophages. The phagocytosis-inducing ability of the autoantibody eluted from senescent cells was absorbed by stored lymphocytes, neutrophils and platelets (Table 1*a*).

Human adult liver cells and human embryonic kidney cells, grown in culture, were also examined for presence of the antigen. The less viable cells which were no longer adherent were collected from culture supernatants by centrifugation (Table 1*b*). Absorption of the IgG eluted from senescent RBCs with both liver and embryonic kidney cells abolished its phagocytosis-inducing ability (Table 1*b*).

As the results of the phagocytosis inhibition assay experiments indicated that the antigen that appears on senescent RBCs also emerges on other somatic cells, I attempted to isolate the antigen from lymphocytes. Sialoglycoprotein mixtures were prepared from lymphocyte membranes in the same manner as described for RBCs (for details, see Fig. 3). Lymphocyte sialoglycoprotein mixtures were processed with the senescent RBC IgG affinity column. Gel electrophoresis of material eluted from the column with glycine–HCl buffer revealed a band migrating at a M_r of ≃62,000, at the same position as the antigen isolated from senescent RBC (Fig. 3). This finding confirms the results obtained with the phagocytosis inhibition assay, indicating that the antigen which appears on senescent RBC also appears on other somatic cells.

Emergence of the ≃62,000-M_r protein initiates binding of IgG autoantibodies *in situ* and phagocytosis of senescent cells by mononuclear phagocytes. This antigen is present on stored lymphocytes, platelets, and neutrophils, and on cultured liver and kidney cells. Therefore, the immunological mechanism[2] for removing senescent and damaged red cells may be a general physiological process for removing somatic cells, at least in mammals.

I thank Drs John Cook, Takashi Makinodan and Andrew Saxon for reviewing the manuscript, and Dr Moses Schanfield for advice and for providing blood-typing reagents, typing donor RBCs and screening donors for anti-RBC antibodies. This work was supported by a NIH grant. Workspace was provided by the Research Service, Wadsworth Medical Center.

Received 7 July; accepted 24 November 1980.

1. Kay, M. M. B. *J. supramolec. Struct.* **9,** 555–567 (1979).
2. Kay, M. M. B. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3521–3525 (1975).
3. Berken, A. & Benacerraf, B. *J. exp. Med.* **123,** 119–144 (1966).
4. Lutz, H. U. & Kay, M. M. B. *Mech. Ageing Dev.* (in the press).
5. Murphy, J. R. *J. Lab. clin. Med.* **82,** 334–241 (1973).
6. Kochwa, S. & Rosenfield, R. *J. Immun.* **92,** 682–692 (1964).
7. Marchesi, V. T. & Andrews, E. P. *Science* **174,** 1247–1248 (1971).
8. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
9. Bennett, G. D. & Kay, M. M. B. *Exp. Hematol.* **9** (in the press).
10. Eckhardt, C. E., Hayes, C. E. & Goldstein, I. J. *Analyt. Biochem.* **73,** 192–197 (1976).
11. Lutz, H., von Däniken, A., Semenza, G. & Bachi, T. *Biochem. biophys. Acta* **552,** 262–279 (1979).
12. Hollister, P. J., Bennett, G. D., Onari, K., Kay, M. M. B. & Makinodan, T. *Gerontologist* (in the press).
13. Silverberg, M., Furthmayr, H. & Marchesi, V. T. *Biochemistry* **15,** 1448–1454 (1976).
14. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **10,** 2606–2617 (1971).
15. Dodge, J. T., Mitchell, C. & Hanahan, D. *Archs Biochem. Biophys.* **100,** 119–130 (1963).

Self-MHC-restricted cytotoxic T-cell response without thymic influence

Istvan Ando

Institute of Genetics, Biological Research Center of the
Hungarian Academy of Sciences, H-6701, Szeged, PO Box 521,
Hungary

Mikko Hurme

Department of Bacteriology and Immunology, University of
Helsinki, Haarmaninkatu 3, SF-00290 Helsinki 29, Finland

The thymus has a central role in directing the maturation and differentiation of T cells[1]. These processes involve movement of cells of haematopoietic origin through the thymus and are supposed to require intimate physical contact between T lymphocytes and cells of the thymic epithelium. Cytotoxic T cells (CTLs) induced by virus-infected or hapten-modified autologous cells exhibit dual specificity: for the products of the self major histocompatibility complex (MHC, H–2 in the mouse) and for the foreign antigen[2–5]. It has been proposed that 'self' specificity of MHC-restricted CTLs is determined by the MHC-encoded antigens expressed on the surface of thymic epithelial cells[6]. We show here that the differentiation of MHC-restricted CTLs can take place without thymic influence. In these experiments, self-MHC-restricted CTLs differentiate from spleen cells of *nu/nu* mice which lack thymuses after stimulation with 2,4,6-trinitrophenyl (TNP)-coupled autologous cells in the presence of T-cell growth factor.

Spleen cells of C57BL/6 *nu*/+ (*nu*/+) (H–2^b) and C57BL/6 *nu/nu* (*nu/nu*) mice were stimulated with TNP-coupled autologous cells *in vitro*. Cells were collected 5 days after stimulation and tested for lytic activity and specificity. An H–2-restricted cytotoxic T-cell response was generated from the cells of *nu*/+ mice, whereas *nu/nu* mice did not respond to the antigen (Table 1). However, when T-cell growth factor (TCGF) (Table 1) was added to the cultures, *nu/nu* cells responded to TNP-derivatized autologous cells. The cytotoxic cells killed preferentially TNP-derivatized autologous cells; TNP-coupled allogeneic cells (B10.BR, H–2^k) were killed to a lesser extent. The level of the cross-reaction was comparable with the cross-reaction by effector cells of *nu*/+ origin. The specificity of the population reacting with TNP-coupled autologous cells was further characterized by nonradiolabelled target-cell inhibition. The results presented in Fig. 1*a* show that the population exhibits a strict specificity. The specificity is restricted to both the self-MHC antigens and to the hapten: autologous cells derivatized with 4-hydroxy-5-iodo-3-nitrophenyl (NIP) or allogeneic cells derivatized with TNP did not compete with TNP-derivatized autologous target cells. Cytotoxic T cells generated in secondary cultures showed the same specificity (Fig. 1*b*). The *nu/nu* effector cells lysing TNP-modified autologous target cells specifically were characterized with respect to the expression of Thy-1 antigen in complement-dependent cytotoxic assay by using an antibody to Thy-1.2 antigen (Table 2). Anti-Thy-1.2 antibody and complement treatment of the effector-cell populations abolished the killer activity demonstrating that these cytotoxic cells were T lymphocytes.

The existence of cytotoxic cells in *nu/nu* mice has been observed previously[7]. The cytotoxic activity was mediated by natural killer (NK) cells, which are known to be Thy-1 negative; their specificity does not show restriction to MHC and their activity is usually determined in a 24-h cytotoxicity assay[8]. The cytotoxic cells induced by TNP-modified autologous cells in the presence of TCGF show all the characteristics of CTLs. They carry Thy-1 antigen, their specificity is restricted to the self-MHC structures and their activity can be determined in a 4-h cytotoxicity assay. The presence of alloreactive pre-effector T cells has been observed in the peripheral lymphoid organs of *nu/nu* mice[9]. These cells were able to differentiate to alloreactive CTLs in the presence of TCGF. Therefore it is supposed that the *nu/nu* mouse is defective in the production of a factor which helps pre-effector T cells to differentiate to immunocompetent, alloreactive, CTL precursor cells. Our experiments show that CTLs with specificity to TNP-modified self-MHC structures are differentiating from *nu/nu* spleen cells after stimulation with hapten-modified autologous cells in the presence of TCGF.

0028–0836/81/060494—02$01.00

Table 1 The specificity of anti-TNP CTL responses of *nu*/+ and *nu*/*nu* mouse spleen cells

Responder	Stimulator	Target	% Specific ^{51}Cr release at A:T 30:1	10:1	3.3:1
nu/+	—	B6–TNP	−6.2	−3.4	−1.0
nu/+	CM	B6–TNP	4.3	−1.2	−1.8
nu/+	B6–TNP	B6	−1.0	−1.0	−3.5
		B6–TNP	61.9	44.6	21.9
		B10 · BR	9.3	3.2	0.6
		B10 · BR–TNP	22.3	8.9	1.6
nu/*nu*	—	B6–TNP	−3.6	−2.2	0.6
	CM	B6–TNP	−10.8	−4.2	−1.7
	B6–TNP	B6–TNP	−4.8	−4.1	−3.0
	B6–TNP + CM	B6	−3.9	−2.3	−0.3
		B6–TNP	38.5	11.4	5.6
		B10 · BR	3.5	2.1	0.0
		B10 · BR–TNP	15.5	0.5	0.6

The generation and assay of CTLs of *nu*/+ origin were carried out as described previously[4]. Briefly, 5×10^7 responding cells and 2×10^7 TNP-derivatized[12] mitomycin C-treated stimulator cells were co-cultured in 25-cm^2 tissue culture flasks in RPMI 1640 medium containing 10% fetal calf serum (Gibco), supplemented with 10 mM HEPES, 200 mM glutamine, 50 IU ml^{-1} penicillin, 50 μg ml^{-1} streptomycin and 5×10^{-5} M 2-mercaptoethanol. After 5 d of co-culture in a humidified CO_2 incubator, the cells were collected and their cytotoxic activity and specificity determined. The target cells were concanavalin A (Con A)-stimulated T-cell blasts labelled with ^{51}Cr ($Na_2{}^{51}CrO_4$; Radiochemical Centre). The attacker (A) and target (T) cells were incubated at different ratios for 4 h, then the supernatants were collected for γ counting. The corrected per cent lysis was calculated according to the following formula:

$$\frac{\text{experimental release} - \text{medium release}}{\text{maximal release} - \text{medium release}}\times 100$$

Medium release was 12–34% of the maximal release. CTLs from *nu*/*nu* spleens were generated in the same way, except that conditioned medium (CM), containing TCGF, was added to the cultures. The CM was prepared according to the method of Gillis *et al.*[13], using rat spleen cells (strain SDK).

Table 2 Demonstration of the presence of Thy-1 antigen on cytotoxic cells derived from the spleens of *nu*/*nu* mice

CTL population	Treatment	% Specific ^{51}Cr release
Primary	Medium	20.8
	Anti-Thy-1,2	26.8
	Anti-Thy-1,2 + C′	0.9
	C′	20.1
Secondary	Medium	31.8
	Anti-Thy-1,2	41.3
	Anti-Thy-1,2 + C′	0.0
	C′	33.7

The antibody-plus-complement (C′) treatment of the attacker cells was carried out in the wells of U-bottomed microtitration plates. Attacker to target cell ratio was 20 : 1 for primary and 5 : 1 for secondary effector cells. Target cells were TNP-derivatized Con A blasts of B6 mice. Attacker cells were incubated with a monoclonal anti-Thy-1.2 antibody (supernatant F7D5) at the predetermined optimal dilution (1 : 50) for 30 min at 24 °C in RPMI 1640 medium. After incubation with the antibody, the cells were washed twice with medium and rabbit serum was added to the cells as a source of complement (1 : 8 dilution). The cells were further incubated for 40 min at 37 °C, then washed twice. 10^4 target cells were added per well and incubated for another 4 h at 37 °C in a CO_2 incubator, then the supernatants were collected for γ counting.

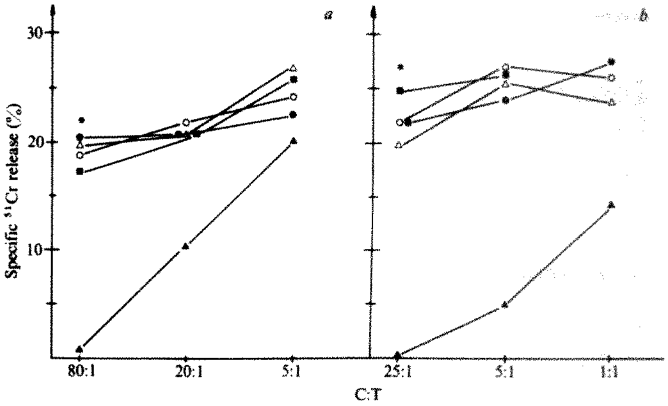


Fig. 1 Specificity of primary (*a*) and secondary (*b*) CTL populations of *nu*/*nu* mice reacting with TNP-derivatized autologous target cells. Secondary CTLs were generated by restimulating 10^5 cells 10 d after *in vitro* priming with 10^5 TNP-derivatized syngeneic lymphocytes in the presence of TCGF. The cells were collected 4 d after restimulation and their cytotoxic activity and specificity were determined. Target cells were TNP-derivatized Con A-induced T-cell blasts of B6 mice. Nonradiolabelled competitor cells were spleen cells, their modification with TNP was carried out as described in Table 1 legend. NIP derivatization was as described previously[14]. Attacker cells were seeded in 50-μl volume in the wells of U-bottomed microtitre plates, in triplicates. 50 μl of nonradiolabelled competitor (C) and 100 μl radiolabelled target (T) cells were added at 4 °C, the cells were spun briefly and incubated for 4 h at 37 °C, and the supernatants collected for γ-counting. The attacker to target cell ratio was 20 : 1 in the primary response and 1 : 1 in the secondary response. Competitors were: C57Bl/6 (△), C57Bl/6–TNP (▲), B10 · BR (○), B10 · BR–TNP (●) or C57Bl/6–NIP (■) cells. * Marks the level of kill in the absence of competitor cells.

Several hypotheses have been proposed to explain the thymic function in determining the H-2 restriction specificity of CTLs. These can be divided into three categories: (1) the thymus drives differentiation, that is, the high mutation rate of germ-line genes within the thymus, creates the specificities; (2) the thymus selects cells for further maturation from an already differentiated repertoire; (3) the thymus only influences the maturation of T cells. Our data strongly oppose the first alternative.

The possibility that *nu*/*nu* mouse spleen cells would not behave in the same way as prothymocytes from normal mice was also considered. The possible residual thymic function in *nu*/*nu* mice, as well as the maternal (*nu*/+) origin of the precursors, are difficult to exclude. Therefore, we have also used prothymocytes from normal mice (bovine serum albumin gradient purified[10], Thy-1 negative) as responders and the results obtained were similar to those with *nu*/*nu* spleen cells described above. Hünig and Bevan[11] have also reported that the anti-TNP CTL responses of nude mice are H-2 restricted.

I.A. was supported by a scholarship from the Hungarian and Finnish Academies of Sciences. We thank Professor Olli Mäkelä for useful discussions and for critical revision of the manuscript, and Dr E. Simpson for supernatant F7D5.

Received 11 June; accepted 14 November 1980.

1. Stutman, O. *Immun. Rev.* **42,** 138–184 (1978).
2. Zinkernagel, R. M. & Doherty, P. C. *Nature* **251,** 547–548 (1974).
3. Shearer, G. M., Rehn, T. G. & Schmitt-Verhulst, A. M. *Transplantn Rev.* **29,** 222–247 (1976).
4. Gordon, R. D., Simpson, E. & Samuelson, L. E. *J. exp. Med.* **142,** 1108–1120 (1975).
5. Bevan, M. J. *J. exp. Med.* **142,** 1349–1364 (1975).
6. Zinkernagel, R. M. *et al. J. exp. Med.* **147,** 882–896 (1978).
7. Paige, C. J., Figarella, E. F., Cuttito, M. J., Cahan, M. & Stutman, O. *J. Immun.* **121,** 1827–1835 (1978).
8. Herberman, R. B., Nunn, M. E., Holden, H. T. & Lavrin, D. H. *Int. J. Cancer* **16,** 230–239 (1975).
9. Gillis, S., Union, N. A., Baker, P. E. & Smith, K. A. *J. exp. Med.* **149,** 1460–1476 (1979).
10. Scheid, M. P., Goldstein, G. & Boyse, E. A. *J. exp. Med.* **147,** 1727–1743 (1978).
11. Hünig, T. & Bevan, M. J. *J. exp. Med.* **152,** 688–702 (1980).
12. Shearer, G. M. *Eur. J. Immun.* **4,** 527–533 (1974).
13. Gillis, S., Ferm, M. M., Du, W. & Smith, K. A. *J. Immun.* **120,** 2027–2032 (1978).
14. Andó, I. & Kisielow, P. *Eur. J. Immun.* **9,** 211–213 (1979).

Interferon-neutralizing antibodies in a patient treated with human fibroblast interferon

Angelika Vallbracht*†, J. Treuner‡, B. Flehmig*, K.-E. Joester§ & D. Niethammer‡

* Department for Medical Virology and Epidemiology of Virus Diseases, Hygiene Institute, University of Tübingen, D-7400 Tübingen 1, FRG
‡ Division of Hematology, Department of Pediatrics, University of Tübingen, D-7400 Tübingen 1, FRG
§ Dr Rentschler & Co., Laupheim, FRG

During recent years clinical trials have shown that human leukocyte interferon (HuIFN-α) may be useful in the treatment of cancer, but very little has been done concerning the possible use of human fibroblast interferon (HuIFN-β). Treuner *et al.*[1] recently reported the successful treatment of a nasopharyngeal carcinoma with HuIFN-β: in the course of IFN-therapy a HuIFN-β neutralizing activity appeared in the serum of this patient. We report here that such activity is due to IgG antibodies—this study is the first to present evidence for antigenicity of IFN in a homologous system.

The IFN used in therapy was a diploid cell-strain IFN (HuIFN-β) (Rentschler Co., Laupheim) produced in human foreskin fibroblasts (FS-4) by superinduction with poly(I · C), cycloheximide and actinomycin D. The IFN preparations were purified and concentrated using perchloric acid and ammonium sulphate precipitation followed by column chromatography and ultrafiltration. The specific activity of the HuIFN-β preparation used was 1–2 × 10⁶ IU per mg protein.

The patient S.R. was treated by i.v. infusions with 4 × 10⁶ units of interferon daily for 5 weeks, followed by three applications per week[1]. Repeated analyses over a period of 6 months showed measurable IFN serum levels after injection[1]. Subsequent control analyses failed to show IFN activity after i.v. injection of 4 × 10⁶ IU. Therapy was stopped after treatment for 10 months (using a total of ~5 × 10⁸ units). The serum was analysed for neutralizing activity against HuIFN-β and immunoglobulins were purified. Unfortunately, no earlier serum samples were available.

Serial 2-fold dilutions of serum were incubated for 1 h at 37 °C and 2 h at 4 °C with eight units of either FS-4 HuIFN-β or HuIFN-α. IFN assays were done by plaque reduction assay in RITA cell culture microplates (monkey kidney cell line) using vesicular stomatitis virus (VSV) as challenge. A simultaneous titration of the IFN used was included in each case and the lowest dilution of antiserum examined for direct inhibition of VSV. IFN titres were expressed in terms of HuIFN-β reference standard G-023-902-527 (from NIH, Bethesda), and the neutralizing titre was taken as the highest dilution of serum that inhibited by 50% the protective activity of at least eight units of IFN. The anti-HuIFN-β titre of untreated serum from S.R. was about 10^{-3} using HuIFN-β (Rentschler) or HuIFN-β (NIH) as antigen, or when HuIFN-β (Rentschler) in the serum of an IFN-treated child was used. A neutralization assay in which the concentration of HuIFN-β was varied showed that 50 µl of serum from patient S.R. were able to neutralize ~10⁴ units of HuIFN-β. Substitution of human serum for fetal calf serum did not induce a change in the anti-HuIFN-β titre. Preincubation of RITA cells with serum from patient S.R. for 20 h inhibited the activity of eight units of HuIFN-β up to a dilution of 1:8—this inhibition could be due to HuIFN-β receptor antibodies or cross-reacting antigens of RITA and FS-4 cells.

Attempts to show neutralizing activity against HuIFN-α in the serum from patient S.R. were unsuccessful[2]. The assay was performed with eight units of HuIFN-α mixed with an appropriate amount of FS-4 cell preparation to rule out the possibility that the presence of antibodies to antigens of the FS-4 cell preparation other than HuIFN-β co-precipitated with IFN to produce a reaction which was erroneously interpreted as HuIFN-β neutralization. No inhibitory effect of serum from patient S.R. against HuIFN-α was seen in these conditions. However, we cannot exclude the possiblity that the antibody observed was produced against a complex of HuIFN-β + FS-4 protein, which could be a result of the method of purification.

Serum taken from the patient S.R. 9 months before IFN therapy showed no neutralizing activity against HuIFN-β. As we had no earlier preserum samples we were unable to check the possibility that disturbed immune function might have resulted in production of autoantibodies against endogeneous HuIFN-β.

Serum taken from patient S.R. after IFW therapy was heated at 56 °C for 30 min before use then 1 ml was absorbed three times with a total of 6 × 10⁸ FS-4 cells; serum–cell mixtures were incubated with agitation at 37 °C for 30 min and 4 °C for 30 min. Serum was centrifuged at 6,000*g* for 30 min to remove cells then the total IgG content measured by immunodiffusion with Tripartigen plates (Behring, Marburg)—this was reduced from

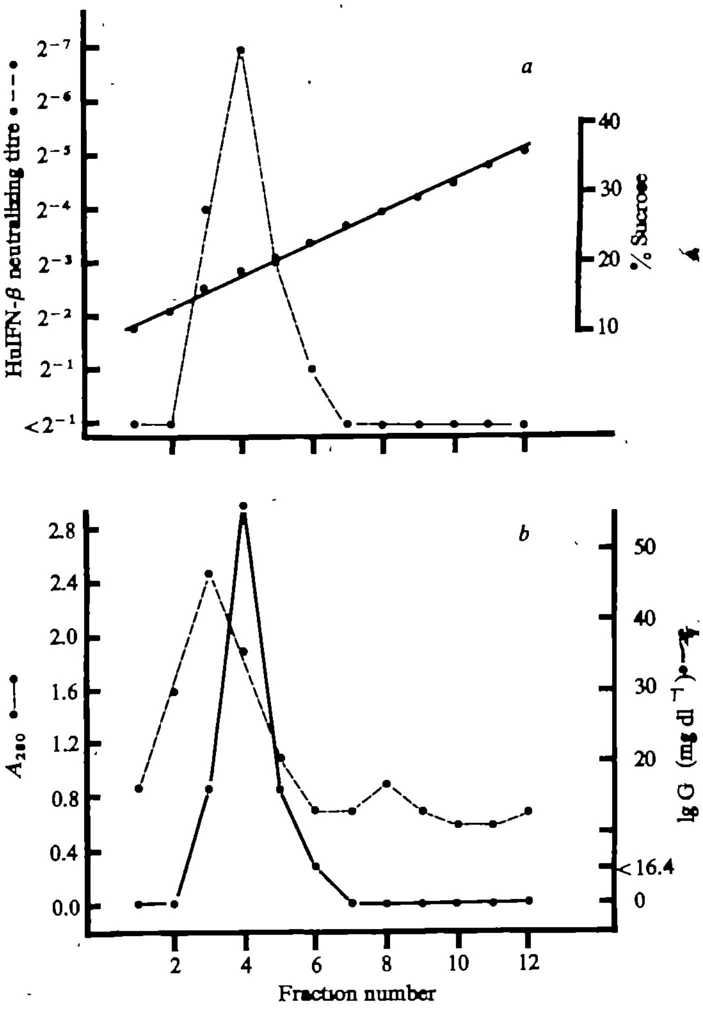


Fig. 1 *a*, Determination of HuIFN-β neutralizing titre. *b*, Total protein content and concentration of IgG after sucrose density-gradient centrifugation of FS-4 absorbed serum from the patient S.R. Serum was centrifuged in a 10–37% sucrose density-gradient for 20 h using a SW-40 rotor (Beckman Instruments, Palo Alto, California) at 154,000*g*. Fraction numbering was started at the top of the gradient. Determination of anti-HuIFN-β titre (broken line, *a*), IgG concentration (solid line, *b*) and total protein content (A_{280}, broken line, *b*) was done with fractions of the same gradient.

---

† To whom correspondence should be addressed

900 mg dl^{-1} before absorption to 550 mg dl^{-1} without diminishing the anti-IFN titre. Preincubation of RITA cells with the absorbed serum did not inhibit HuIFN-β activity.

To demonstrate that the HuIFN-β neutralizing activity was in fact due to antibodies, activity was determined after sucrose density-gradient centrifugation of FS-4 absorbed serum from S.R. Total protein content (E_{280}) and concentrations of IgG in sucrose density-gradient fractions of serum are shown in comparison with anti-HuIFN-β titres in Fig. 1. The peak of anti-HuIFN-β activity was found in fraction 4 (Fig. 1*a*) which correlates with the IgG (Fig. 1*b*) and the IgA peak (not shown). The IgM peak (fraction 8) showed no measurable HuIFN-β neutralizing activity. In addition, serum IgG from patient S.R. was purified by ion-exchange chromatography to demonstrate that the anti-HuIFN-β activity was due to IgG antibodies. FS-4 cell-absorbed serum was precipitated with saturated ammonium sulphate, *p*H 6.5, and the precipitate dialysed against 0.02 M phosphate buffer, *p*H 8.0. IgG was prepared from the precipitate by chromatography on DEAE-Sephacel column (Pharmacia, Uppsala, Sweden) equilibrated with the same buffer. The protein peak passed through with the first buffer volume contained most of the anti-HuIFN-β activity. Immunodiffusion showed that there was no IgA or IgM contamination of the purified IgG. Immunoelectrophoreses of 0.7 mg IgG ml^{-1} against rabbit anti-human serum protein (Medac, Hamburg) revealed only one component in the IgG preparation.

The first successful demonstration of antiserum to an IFN was reported by Paucker and Cantell[3]. However, studies on the antigenicity of IFN have shown that they are not antigenic in homologous systems[4]. In a number of clinical trials no neutralizing activity could be demonstrated in the sera of any patients, even after treatment with several millions of human IFN units[5,6].

This study presents a strong case for the presence of antibodies to HuIFN-β in the serum of patient S.R. treated successfully with $\sim 5 \times 10^8$ HuIFN-β units. A number of other patients in the Department of Pediatrics in Tübingen were treated with the same preparation of HuIFN-β—three of these had a nasopharyngeal carcinoma[7]. Total remission of the carcinoma after HuIFN-β therapy was shown only in S.R. who was the only patient in which HuIFN-β neutralizing antibodies could be demonstrated, although several others were treated with equally high doses.

We are now attempting to determine whether the patient's cells lack the genes for HuIFN-β and further experiments are being done to clarify this phenomenon of anti-IFN production and its possible clinical significance.

Received 26 September; accepted 27 November 1980.

1. Treuner, J. *et al. Lancet* **i,** 817–818 (1980).
2. Berg, K., Ogburn, C. A., Paucker, K., Mogensen, K. E. & Cantell, K. *J. Immun.* **114,** 640–644 (1975).
3. Paucker, K. & Cantell, K. *Virology* **18,** 154–147 (1962).
4. Levy-Koenig, R. E., Mundy, M. J. & Paucker, K. *J. Immun.* **104,** 785–790 (1970).
5. Strander, H., Cantell, K., Carlstrom, G. & Jakobson, P. A. *J. natn. Cancer Inst.* **51,** 733–739 (1973).
6. Desmyter, J. *et al. Lancet* **ii,** 645–648 (1976).
7. Treuner, J. *et al. J. Interferon Res.* (submitted).

Functional reactivation of the deafferented neostriatum by nigral transplants

A. Björklund & U. Stenevi

Department of Histology, University of Lund,
S-223 62 Lund, Sweden

S. B. Dunnett & S. D. Iversen

Department of Experimental Psychology, University of Cambridge,
Cambridge CB2 3EB, UK

Functional deficits following brain lesions can be due not only to the disruption of conduction in specific input and output pathways passing through the site of injury, but also to the loss of important regulatory systems controlling the functional state of neuronal circuitries in areas distant from the lesion. For example, the behavioural disturbances that result from lesions of the nigrostriatal dopamine (DA) pathways[1-3] can be reversed by administration of dopamine receptor-activating drugs, such as L-dopa or apomorphine[4-6]. This suggests that the lesioned dopaminergic system, rather than conveying specific input and output signals, is normally acting on neuronal machineries whose activity levels are set by the activity at the dopaminergic synapses. Thus the neurological deficits resulting from these lesions are due to functional inactivation of otherwise intact neostriatal circuitries. Previous studies have shown that intracerebral transplants of embryonic substantia nigra can compensate for drug-induced[7-9] as well as spontaneous asymmetric motor behaviour[11] (expressed as a tendency to move in circles towards the lesioned side), whereas the sensorimotor asymmetry, which is pronounced in rats with a unilateral lesion of the nigrostriatal DA pathway[4,10], was unaffected by the transplant[11]. We report here that restoration of striatal dopaminergic neurotransmission by nigral transplants in animals with bilateral, complete lesions of the nigrostriatal DA pathways can reinstate not only certain aspects of spontaneous motor behaviour, but also sensorimotor orientation and sensory attention on the side of the body contralateral to the graft.

Young female Sprague–Dawley rats (180–200 g at the time of surgery) initially received a unilateral right lesion of the nigrostriatal DA pathway, produced by an intracerebral injection of 8 µg 6-hydroxydopamine (6-OHDA; Hässle) into the pathway[12]. The completeness of the lesion was tested by assessing the intensity of the amphetamine-induced rotational response according to Ungerstedt and Arbuthnott[13]. When injected intraperitoneally with 5 mg per kg metamphetamine 1 week after the 6-OHDA lesion, all the rats included in the study showed a mean of >8 complete 360° turns per min. This turning rate has previously been found to correspond to a reduction in striatal DA of 97% or more[9].

About 4 weeks after the 6-OHDA lesion the rats received transplants of the substantia nigra region taken from 16–17-day rat embryos (crown to rump length 15–20 mm) as described elsewhere[7,9]. The transplants were placed on the dorsal surface of the denervated neostriatum, within a suction cavity made through the anterior parietal cortex and the corpus callosum. Eight control rats received the cavity lesion but no transplants. Six months after transplantation, 14 of the transplanted rats which showed complete compensation of their initial amphetamine-induced turning response (Table 1) were selected for further behavioural testing. (The compensation of amphetamine turning has previously been shown to be a good criterion for animals with surviving transplants and with significant dopaminergic re-innervation of the dorsal part of the denervated caudate-putamen[9]. This was confirmed in the final fluorescence histochemical analysis which demonstrated DA-rich surviving transplants and significant DA fibre ingrowth in each case.)

Seven months after the first 6-OHDA lesion, all 22 rats received a second 6-OHDA lesion to the contralateral nigrostriatal DA pathway. They were maintained by intragastric tube-feeding for 8 weeks, giving 25–40 ml per day of a high-energy liquid diet preparation (Vivonex, Vitrum). All surviving rats were finally killed and the brains taken for fluorescence microscopical analysis using the ALFA method of Lorén *et al.*[14]. The rats underwent behavioural testing before and after the second lesion, which involved observations on amphetamine-induced and spontaneous asymmetric motor behaviour, arm

0028-0836/81/060497—03$01.00

Table 1 Behavioural recovery in nigra-transplanted rats

	Unilateral (right) lesion		Bilateral lesion	
	No transplant (*n* = 8)	Right transplant (*n* = 14)	No transplant (*n* = 8)	Right transplant (*n* = 14)
Amphetamine-induced rotation (net ipsilateral turns per min)				
Before transplantation (1 week after 1st lesion)	14.0 ± 2.5	14.0 ± 0.7*	—	—
3 months after 1st lesion	16.3 ± 1.4	2.7 ± 1.0†	—	—
2 months after 2nd lesion	—	—	0.9 ± 1.4	−11.6 ± 0.8†
Spontaneous rotation (probability of left turn)				
Without activation	0.17 ± 0.04	0.32 ± 0.05†	0.57 ± 0.12	0.92 ± 0.04†
Tail-pinch activation	0.24 ± 0.06	0.46 ± 0.05†	0.61 ± 0.11	0.91 ± 0.03†
Arm preference in T-maze (probability of left choice)	0.03 ± 0.02	0.32 ± 0.07†	0.41 ± 0.09	0.79 ± 0.06†
Sensorimotor asymmetry (right–left scores)				
Without activation	3.00 ± 0.82	4.00 ± 0.72*	0.29 ± 0.52	−4.33 ± 0.67†
Tail-pinch activation	4.17 ± 1.05	3.23 ± 0.63*	1.14 ± 0.77	−4.67 ± 0.50†

Amphetamine-induced rotation was recorded in plastic bowls over 60 min after intraperitoneal administration of 5 mg per kg metamphetamine, after the method of Ungerstedt and Arbuthnott[13]. Scores are net ipsilateral (right–left) turns per min. Spontaneous rotation was recorded in the same bowls, and quarter-turns in each direction were recorded over a 3-min period. Ten min later, a paper clip was attached to the rat's tail to provide mild activation, and again quarter-turns were recorded over a 3-min period. Maze testing: rats were tested for arm preference in an unbaited T-maze of conventional design, with and without tail-pinch activation, as described in detail elsewhere[11]. Each test consisted of two trials separated by a 5–10-min interval. Three tests were carried out, each separated by several days, and the scores for each animal combined as no differences were found between tests. Sensorimotor asymmetry was tested in a battery modified from Marshall and Teitelbaum[10,20], and described in detail elsewhere[11]. Briefly, the orientation to a light pin-prick at different points on the body surface, to touch of the whiskers and snout, and to the odour of an ammonia-soaked swab, were each rated on a three-point scale (0 = absent, 1 = weak, 2 = normal). The asymmetry scores are the difference between total scores on the two sides of the body (right–left). All tests were performed 3–4 months after the initial unilateral lesion and transplantation in the unilateral animals. They all received bilateral lesions at 7 months, and were retested 6–8 weeks later. The amphetamine rotation and sensorimotor tests indicate net (right–left) scores, with positive values denoting bias towards the right (transplant/initial lesion) side. The spontaneous rotation and maze tests indicate the probability of a left turn, scores <0.5 denote a right bias, and >0.5 a left bias. All values are means ±s.e.m. and comparisons by unrelated Student's *t*-test.

* Not significant, $P>0.01$

† $P<0.001$.

preference in a T-maze, and asymmetry in sensorimotor responding, as summarized in Table 1.

After the second 6-OHDA lesion both the controls and the transplanted rats became severely akinetic. Both groups of rats also became aphagic and adipsic, and they had to be maintained by intragastric tube-feeding throughout the 8-week test period.

The akinetic state of the animals was partially reversed by tail-pinch activation. When activated in this way, the non-transplanted control rats showed no side bias in their spontaneous turning behaviour and no asymmetric posture. In contrast, the transplanted rats showed a high degree of turning away from the transplant side (towards the left) (right-hand panels in Table 1) and asymmetric posture, with the head, body and tail describing a curve concave on the side of the second lesion. Without tail-pinch activation the transplanted rats showed an equally pronounced bias for direction of turning in the spontaneous rotation test, although the turning rate was much lower. Before the second lesion (left panels in Table 1), the lesioned, non-transplanted controls showed a strong rotational bias towards the lesioned side; this bias was significantly attenuated in the transplanted rats, both with and without tail-pinch activation.

When tested for arm preference in the T-maze, which can be taken as a measure of the relative degree of sensory attention on the two sides of the rat's body, before the second lesion the non-transplanted controls showed a strong bias (97%) for the right arm contralateral to the intact nigrostriatal pathway. This bias was significantly reduced in the transplanted rats. After the second lesion the transplanted rats showed a marked preference (79%) for the left arm, contralateral to the transplant, whereas the bilaterally lesioned controls showed no such asymmetry (Table 1).

When performed before the second lesion, the sensorimotor orientation test demonstrated an equally marked sensory inattention on the left side of the body in both the transplanted and non-transplanted rats. This deficit was unaffected by mild tail-pinch activation. After the second lesion the controls became equally unresponsive on both sides of the body, whereas the transplanted rats showed a marked asymmetry, with significantly greater response towards stimuli applied on the left side of the body, contralateral to the transplant (Table 1). Thus, with respect to both motor asymmetry, choice behaviour in the T-maze and sensorimotor orientation, the bilaterally lesioned rats with unilateral nigral transplants showed a performance similar to that of rats with a unilateral lesion of the nigrostriatal system.

Thus, our results support the conclusion that an ectopic, intracortical nigral transplant, without any external (that is, drug-induced) activation, can induce recovery in certain aspects of motor behaviour, and in sensory attention and sensorimotor orientation specifically on the side contralateral to the transplant, in the spontaneously behaving animal. In contrast, the severe akinesia, aphagia and adipsia of the bilaterally lesioned rats were unaffected.

Figure 1 shows in diagrammatic form a possible model for the mechanism of action of the ectopic nigral transplant, based on the one previously proposed by Stricker and Zigmond[2]. In this model the DA systems regulate the behavioural responsiveness through modulation of an inhibitory control mechanism. In the intact animal the nigrostriatal DA pathway seems to act as a tonic regulatory system which sets the level of activity in the neostriatal machinery. Removal of this dopaminergic regulatory control function results in inhibition of neostriatal function. Reinstatement of dopaminergic neurotransmission by drugs[4–6], or—as in our experiments—by transplantation of DA neurones, can thus be interpreted as a reactivation of an inhibited, but otherwise intact, neuronal machinery. If we assume that the nigrostriatal DA system is inhibitory in nature (see ref. 15 for discussion), then the likely action of the nigral transplant would be to disinhibit elements of the nigrostriatal circuitry.

In the model in Fig. 1 the nigral transplant acts by reducing the threshold for behavioural response (R) to diverse sensory inputs (I). An interesting implication of this model is that recovery of responsiveness to sensory stimuli in the transplanted rats can occur without any access of sensory information to the transplant. In the intact animal various types of sensory stimuli have been shown to activate the nigrostriatal DA neurones[16,17]. However, in the transplanted animal such activation would not be necessary for behavioural recovery provided that the transplanted neurones have a sufficiently high spontaneous activity. Other parallel observations indicate that this indeed is the case. First, in biochemical determinations of DA and its metabolite, dihydroxyphenylacetic acid[18], the turnover of DA in both the transplant and the re-innervated neostriatum has been found to be as high as that of the intact nigrostriatal system. Second, ^{14}C-2-deoxy-D-glucose autoradiography has revealed a significant glucose uptake in the nigral transplants, amounting to about two-thirds of that of the intact substantia nigra[18], and third, measurements of the rotational response to the DA receptor agonist, apomorphine, in nigra-transplanted rats indicate that re-innervation of the neostriatum by nigral transplants is accompanied by a partial reduction in the supersensitivity of the neostriatal DA receptors[11]. This observation is compatible with a partial restoration of dopaminergic neurotransmission in the re-innervated neostriatum.

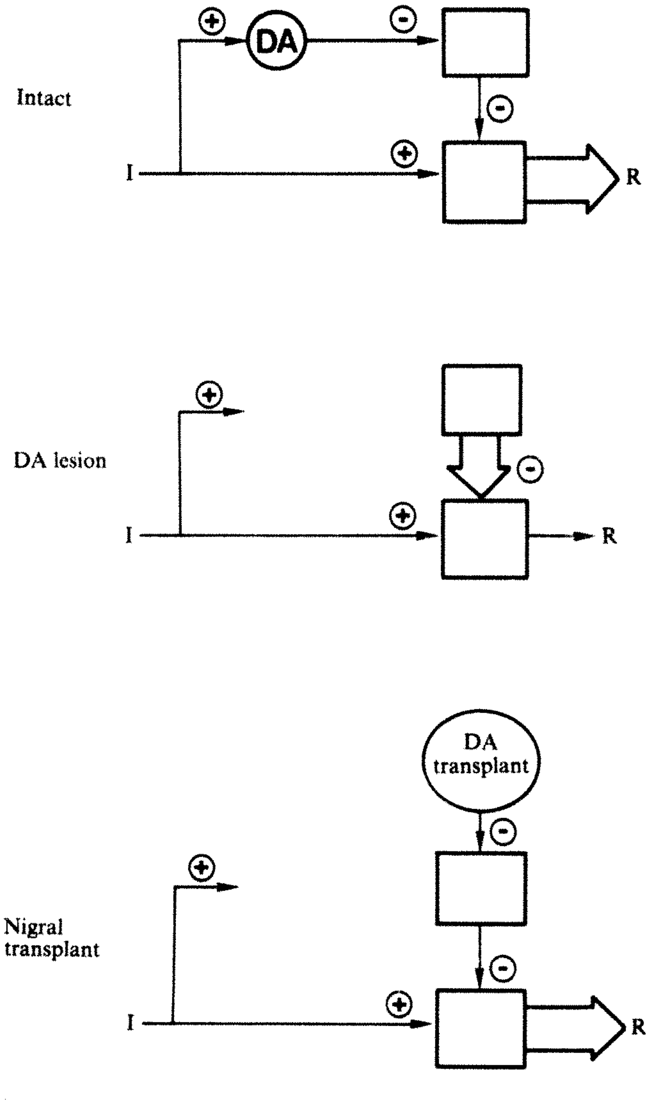


Fig. 1 Proposed mode of action of the DA-producing intracortical nigral transplants in the promotion of behavioural recovery in rats with unilateral or bilateral lesions of the nigrostriatal DA pathway. The behavioural response, R, evoked by an activating sensory input, I, is modulated by the transplant through an inhibitory control mechanism.

An important feature of the behavioural effects of the nigral transplants was the specificity of the transplant-induced recovery. Thus, in the bilaterally lesioned rats the transplants did not induce a diffuse, partial recovery in all aspects of the DA deficiency syndrome. Rather, the recovery appeared discrete and complete, such that in those behaviours where recovery did occur (motor asymmetry and choice behaviour) the transplanted animals performed as well as an intact rat, whereas in those behaviours where recovery did not occur (akinesia, aphagia and adipsia) they performed as badly as the lesioned control rats. This can be explained if one assumes that different parts of the striatal complex subserve, at least partly, different types of function (see ref. 19). If so, the transplant-induced recovery would depend on which parts of the striatal complex are re-innervated from the transplant. With the present intracortical transplantation site the ingrowing DA fibres are confined to the dorsal parts of the head of the caudate-putamen, whereas the ventral and lateral parts, as well as nucleus accumbens, are left denervated. The possibility that this regional restriction of the transplant-induced dopaminergic re-innervation determines the pattern and specificity of behavioural recovery is now being investigated.

These results demonstrate that intracerebral transplants can be functional not only after pharmacological activation but also in the spontaneously behaving animal, and that ectopic nigral transplants, which probably lack their normal afferent inputs, can nevertheless compensate for the loss of the rat's own substantia nigra, at least in certain types of behaviour. Functional reactivation of denervated central nervous system circuitries by neural transplants should thus offer an interesting approach to the analysis of mechanisms of neurological and behavioural recovery after brain or spinal cord damage as well as in animal models of neurodegenerative disorders, such as Parkinson's disease.

We thank Birgit Haraldsson, Ulla Jarl and Gertrude Stridsberg for technical assistance. The work was supported by grants from the Swedish MRC (04X-3874), the Wiberg and Kock Foundations, and the European Training Programme in Brain and Behaviour Research.

Received 11 August; accepted 4 December 1980.

1. Ungerstedt, U. *The Neurosciences, Third Study Program*, 695–977 (MIT Press, 1974).
2. Stricker, E. M. & Zigmond, M. J. *Progress in Psychobiology and Physiological Psychology*, 121–188 (Academic, New York, 1976).
3. Marshall, J. F. & Teitelbaum, P. *Handbook of Psychopharmacology* Vol. 7, 201–229 (Plenum, Oxford, 1977).
4. Ljungberg, T. & Ungerstedt, U. *Physiol. Behav.* **16,** 277–283 (1976).
5. Marshall, J. F. & Ungerstedt, U. *Physiol. Behav.* **17,** 817–822 (1976).
6. Marshall, J. F. & Gotthelf, R. *Expl Neurol.* **65,** 389–411 (1979).
7. Björklund, A. & Stenevi, U. *Brain Res.* **177,** 555–560 (1979).
8. Perlow, M. J. *et al. Science* **204,** 643–647 (1979).
9. Björklund, A., Dunnett, S. B., Stenevi, U., Lewis, M. E. & Iversen, S. D. *Brain Res.* **199,** 307–333 (1980).
10. Marshall, J. F., Richardson, J. S. & Teitelbaum, P. *J. comp. physiol. Psychol.* **87,** 808–830 (1974).
11. Dunnett, S. B., Björklund, A., Stenevi, U. & Iversen, S. D. *Brain Res.* (in the press).
12. Marshall, J. F. & Ungerstedt, U. *Science* **198,** 62–64 (1977).
13. Ungerstedt, U. & Arbuthnott, G. W. *Brain Res.* **24,** 485–493 (1970).
14. Lorén, I., Björklund, A., Falck, B. & Lindvall, O. *J. Neurosci. Meth.* **2,** 277–300 (1980).
15. Moore, R. Y. & Bloom, F. E. *A. Rev. Neurosci.* **1,** 129–169 (1978).
16. Glowinski, J., Niecoullon, A. & Chéramy, A. *Adv. Biochem. Psychopharma.* **19,** 75 (1973).
17. Chiodo, L. A., Antelman, S. M., Caggiula, A. R. & Lineberry, C. G. *Brain Res.* **189,** 544–549 (1980).
18. Ingvar, M., Lindvall, O., Schmidt, R. H., Stenevi, U. & Björklund, A. *Abstr. 4th Eur. Neurosci. Meeting*, Brighton (1980).
19. Dunnett, S. B. & Iversen, S. D. *Behav. Brain Res.* **1,** 497–506 (1980).
20. Marshall, J. & Teitelbaum, P. *J. comp. physiol. Psychol.* **86,** 375–395 (1974).

Central noradrenergic neurones concentrate ^{3}H-oestradiol

Madhabananda Sar & Walter E. Stumpf

Departments of Anatomy and Pharmacology, University of North Carolina, Chapel Hill, North Carolina 27514

Central catecholamines play an important part in the regulation of hormone secretion from the pituitary gland and in the mediation of male and female sexual behaviour[1-6]. Noradrenaline has been shown to stimulate the release of luteinizing hormone (LH) probably by influencing the secretion of LH-releasing hormone (LHRH)[7,8]. The medial preoptic and hypothalamic areas, including the median eminence, contain noradrenaline-containing terminals which originate from discrete noradrenaline-containing cell groups in the lower brain stem. These cell groups have been identified in the pons and the medulla oblongata by histochemical methods and pharmacological experiments[9-14]. Recent studies with antiserum to dopamine-β-hydroxylase (DBH), the enzyme that converts dopamine to noradrenaline, provided evidence for the existence of noradrenergic cells[15,16]. These include the locus coeruleus (group A6), a ventrally located more diffuse but continuous subcoeruleus group (group A5), a cell group located dorsal to the nucleus (n.) dorsalis motorius nervi vagi (group A2) and a cell group in or near the nucleus reticularis lateralis (group A1). Using the thaw-mount autoradiographic technique, oestradiol-concentrating neurones have been localized in many areas of the lower brain stem, including the locus coeruleus, n. tractus solitarii, n. dorsalis motorius nervi vagi, and reticular formation[17] where catecholamine-containing neurones exist. We report here the simultaneous localization, in the same histological section, of ^{3}H-oestradiol and the enzyme dopamine-β-hydroxylase in neurones of the rat lower brain stem with a combined technique of thaw-mount autoradiography and immunohistochemistry[18,19], demonstrating that noradrenaline- or adrenaline-containing neurones are oestradiol target cells.

Ten adult female Holtzman Sprague–Dawley rats, castrated 72 h previously, were injected intravenously with [2,4,6,7-^{3}H]oestradiol-17β, specific activity 95 Ci mmol^{-1}, at a dose of 0.5 μg per 100 g body weight and killed after 1 h. Four ovariectomized rats were used to demonstrate the specificity of localization: two each injected with 500 μg of progesterone or 500 μg of 5-α-dihydrotestosterone, 15 min before ^{3}H-oestradiol. Lower brain stem with pons and medulla oblongata was dissected, placed on a tissue holder and frozen in liquefied propane (−180 °C). Serial frozen sections (4 μm) were cut in a cryostat (Harris), thaw-mounted on photographic emulsion (Kodak NTB3) slides, and exposed in light-proof boxes at −15 °C. The thaw-mount autoradiographic procedure has been described elsewhere[20]. After 8–12 months of exposure, the slides were fixed for 30 s in 3% paraformaldehyde solution (*p*H 7.5) at ice temperature, rinsed briefly in 0.01 M phosphate-buffered saline (PBS) solution at *p*H 7.5, and developed in Kodak D-19 developer (diluted 1:1 in PBS), followed by fixation with Kodak rapid fixer. After photographic processing and subsequent washing with two changes of PBS the autoradiograms were treated with phenylhydrazine–egg albumin to reduce nonspecific staining[19,21]. After a brief wash with PBS, the autoradiograms were incubated at 4 °C for 48 h with a 1:250 or 1:500 dilution of DBH antiserum[22] in PBS containing 1% normal sheep serum and 0.03% Triton X-100. All subsequent steps were carried out at room temperature. After a 3-min wash in PBS, sections were incubated with sheep anti-rabbit γ globulin (1:100 in PBS) for 10 min. After a further 3-min wash with PBS, the autoradiograms were incubated for 20 min with rabbit peroxidase–antiperoxidase complex (1:50). Another 3-min wash in PBS was followed by immersion of the slides for 10 min in 200 ml Tris buffer (0.05 M, *p*H 7.6) containing 150 mg of diaminobenzidine tetrahydrochloride and 15 μl of 30% hydrogen peroxide. Autoradiograms incubated with normal rabbit serum were used to test the specificity of the reaction.

Thaw-mount autoradiograms of pons and medulla oblongata stained immunohistochemically with antiserum to DBH showed DBH-positive cells in certain regions of the hind brain (Figs 1, 2). The strongest DBH staining was observed in cells of the locus coeruleus throughout its extensions, both in its dorsal and ventral divisions. About 30–40% of these cells (group A6) showed nuclear concentration of radioactivity after ^{3}H-oestradiol administration. The localization was specific and probably represented oestradiol because administration of unlabelled oestradiol in excess abolished the nuclear uptake[23], whereas unlabelled progesterone or 5α-dihydrotestosterone did not. A few DBH-positive cells located at the rostral part near the dorso-lateral corner of the fourth ventricle and at the caudal end of the locus coeruleus were not labelled with ^{3}H-oestradiol. In addition to the locus coeruleus, positively stained scattered cells were found around the lateral and dorsal aspects of superior olivary nucleus (group A5) and in regions of the lateral lemniscus (group A7). These cells belong to the lateral tegmental area. About 30% of these DBH-positive cells showed nuclear concentration of ^{3}H-oestradiol. In the ventrolateral medulla oblongata, an accumulation of DBH-positive cells were found in and around the n. lateralis reticularis (group A1)—about 50% of these DBH-positive cells were labelled with ^{3}H-oestradiol. In the dorsomedial medulla oblongata, DBH-positive cells were located in the n. tractus solitarii near the n.

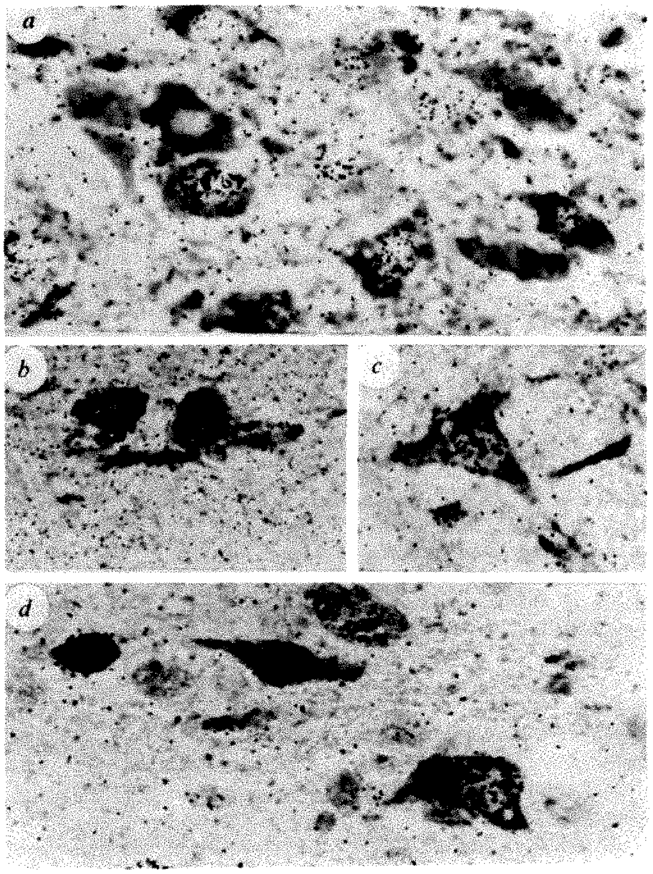


Fig. 1 Thaw-mount autoradiograms of rat lower brain stem 1 h after injection of ^{3}H-oestradiol, stained by immunoperoxidase method using antiserum to DBH (1:500). Concentration of radioactivity is seen in nuclei of DBH-positive cells in *a*, locus coeruleus, *b*, lateral tegmental area, *c*, n. reticularis lateralis, and *d*, n. tractus solitarii. Exposure time, 360 days; 4 μm. ×840 (*a*, *b*, *c*); ×1,050 (*d*).

0028-0836/81/060500—03$01.00

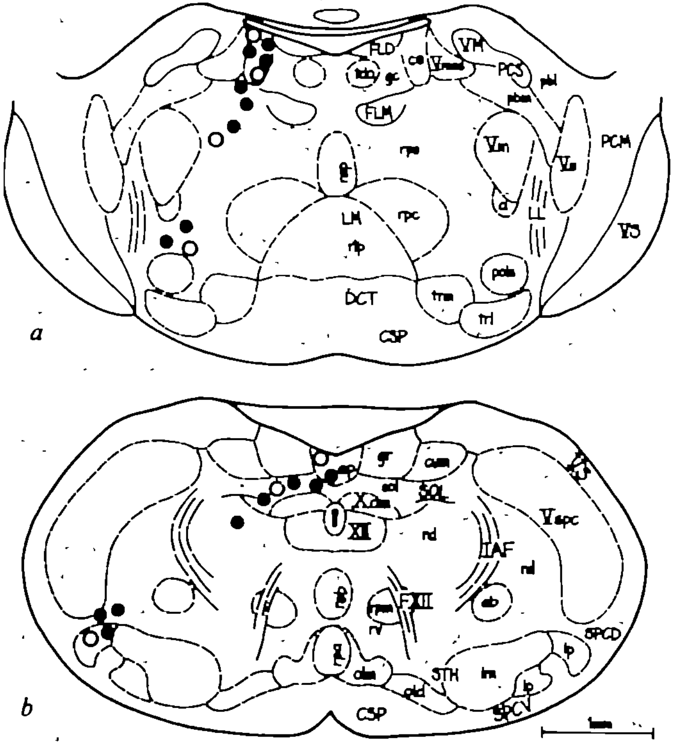


Fig. 2 Schematic drawings of DBH-containing neurones which concentrate ^{3}H-oestradiol (●) in pons (*a*) and in medulla oblongata (*b*). ○, DBH-containing neurones which do not concentrate ^{3}H-oestradiol. Frontal plane. α, nucleus (n.) alpha, ab, n. ambiguus; ap, area postrema; ce, locus coeruleus; CSP, tractus corticospinalis; cum, n. cuneatus medialis, DCT, decussatio corporis trapezoidei; FLD, fasciculus longitudinalis dorsalis, FLM, fasciculus longitudinalis medialis; gc, griseum centrale, gr, n. gracilis; IAF, fibrae arcuatae internae, LL, lemniscus lateralis, LM, lemniscus medialis; lm, n. reticularis lateralis magnocellularis; lp, n. reticularis lateralis parvocellularis; old, n. olivaris accessorius dorsalis; olm, n. olivaris accessorius medialis, pbl, n. parabrachialis lateralis; pbm, n. parabrachialis medialis; PCM, pedunculus cerebellaris medius; PCS, pedunculus cerebellaris superior pols, n. paraolivaris superior, rap, n. raphes pontis; rd, n. reticularis dorsalis medullae oblongata; rpa, n. raphe pallidus, rob, n. raphe obscurus; rpm, n. raphe paramedianus; rpo, n. reticularis pontis oralis; rtp, n. reticularis tegmenti pontis; SOL, tractus solitarius; sol, n. tractus solitarii; SPC, tractus spinocerebellaris; SPCV, tractus spinocerebellaris ventralis, STH, tractus spinothalamicus; tdo, n. tegmentalis dorsalis; trl, n. trapezoides lateralis; trm, n. trapezoides medialis; Vm, n. motorius nervi trigemini; Vmes, n. tractus mesencephali nervi trigemini, VS, tractus spinalis nervi trigemini; Vs, n. sensibilis nervi trigemini, Vspc, n. caudalis tractus spinalis nervi trigemini; Xdm, n. dorsalis motorius nervi vagi, XII, n. nervi hypoglossi.

dorsalis motorius nervi vagi, with a few scattered DBH-positive cells extending caudally into the nucleus commissuralis (group A2): ~70–80% of these positive cells were labelled with ^{3}H-oestradiol. In addition, certain cells of the area postrema showed immunoreactivity to DBH, a few of which were labelled with ^{3}H-oestradiol. When autoradiograms were incubated with the normal rabbit serum and stained immunohistochemically no positive staining was observed. The specificity of DBH antiserum has been established elsewhere[22].

These results provide direct evidence that DBH-positive neurones in lower brain stem are targets for oestradiol and suggest a direct action of oestradiol on the catecholaminergic neurones. This agrees with the previous observation of formaldehyde-induced fluorescence in oestrogen target neurones in rat lower brainstem[24]. The combined formaldehyde-induced fluorescence and autoradiography technique, however, does not distinguish between dopaminergic and noradrenergic neurones. Also, there is the possibility of fading of the fluorophore formed from the catecholamine as a result of exposure time of the autoradiograms. The use of antibodies to transmitter-synthesizing enzyme(s) provides a more specific and sensitive approach for the characterization of catecholaminergic neurones. This combined technique is also of great potential for studying morphological interrelationships between steroid hormones and neuropeptides such as LHRH, somatotropin release inhibiting factor and neurotensin. The positive staining observed here with DBH antiserum does not distinguish between noradrenaline- and adrenaline-producing neurones, because both catecholaminergic neurones contain DBH. Catecholaminergic neurones in regions of the n. tractus solitarii and n. reticularis lateralis may contain both noradrenaline and adrenaline. The question of whether adrenaline-producing neurones are targets for oestradiol can only be clarified by the application of this combination technique using antiserum to phenylethanolamine-*N*-methyltransferase (PNMT), the enzyme specific for adrenaline synthesis[25].

There is considerable evidence for the effects of gonadal steroids on central catecholamine metabolism. After ovariectomy, administration of oestradiol-17β results in changes in hypothalamic noradrenaline turnover rates[26–28], increased tyrosine hydroxylase activity[29] and changes on monoamine oxidase activity[30]. As oestradiol localizes in noradrenergic and, probably, adrenergic neurones, certain effects of gonadal steroids on the regulation of gonadotropin secretion and sexual behaviour are likely to be mediated through direct activation of catecholamine-containing neurones. Results from physiological experiments[31,32], although often conflicting, have favoured this hypothesis. For example, noradrenaline systems mediate the stimulatory effects of oestrogen on the release of prolactin[8,33,34] and of progesterone on release of LH and follicle-stimulating hormone[7,35]. DBH inhibitors have been shown to provoke anti-ovulatory effect, supporting the stimulating role of noradrenergic neurones in the control of LHRH secretion[36]. Thus, the positive oestradiol feedback that produces the pre-ovulatory LH surge probably involves activation of the noradrenaline system by a direct effect on catecholaminergic neurones in the n. tractus solitarii and n. reticularis lateralis of the medulla. Day *et al.*[37] have used the combined localization of retrogradely transported horseradish peroxidase and catecholamine-fluorescence to demonstrate that the noradrenaline innervation of the medial preoptic area is derived from these nuclear groups. Noradrenaline-containing terminals in the medial preoptic area may stimulate LH secretion by influencing the release of LHRH[8,35]. The direct association of noradrenaline-containing terminals with the LHRH-producing cells, however, remains to be demonstrated.

This work was supported by PHS grant NS09914 and NIMH grant HD03110. Antiserum was supplied by T. H. Joh.

Received 24 September, accepted 7 November 1980

1 Kalra, P. S., Kalra, S. P., Krulich, L., Fawcett, C. P. & McCann, S. M. *Endocrinology* **90**, 1168–1176 (1972)
2 Kalra, S. P. & McCann, S. M. *Prog. Brain Res.* **39**, 185 (1973)
3 Fuxe, K., Hökfelt, T. & Jonsson, G. J. *Comp. Physiol. Psychol.* **89**, 556 (1975).
4. Malmnas, C. O. in *Pharmacology, Biochemistry and Behavior* Vol. 4, 521 (Ankho Intl Inc., 1976)
5 McCann, S. M. & Moss, R. L. *Life Sci.* **16**, 833 (1975)
6 Herndon, J. G. Jr *Physiol. Behav.* 17, 143 (1976).
7. Simpkins, J. W. & Kalra, S. P. *Brain Res.* **170**, 475 (1979).
8 Honma, K. & Wuttke, W. *Endocrinology* **106**, 1848 (1980)
9 Dahlström, A. & Fuxe, K. *Acta physiol. scand.* **62**, Suppl. 232, 1, 55 (1964)
10 Andén, N. E. *et al. Acta physiol. scand.* **67**, 313 (1966)
11. Fuxe, K. & Dahlström, A. *Pharmac. Rev.* **18**, 727 (1966)
12 Corrodi, H., Fuxe, K., Hamberger, B. & Ljungdahl, A. *Eur. J. Pharmac.* **12**, 145 (1970)
13. Palkovits, M. & Jacobowitz, D. M. *J. comp. neurol.* **157**, 29 (1974)
14 Moore, R. Y. & Bloom, F. E. *A. Rev. Neurosci.* **2**, 113 (1979).
15 Fuxe, K., Goldstein, M., Hökfelt, T. & Joh, T. H. *Res. Commun. chem. Path. Pharmac.* **1**, 627 (1970)
16. Swanson, L. W. & Hartman, B. K. *J. comp. Neurol.* **163**, 467 (1975)
17. Stumpf, W. E. & Sar, M. in *Anatomical Neuroendocrinology* (eds Stumpf, W. E. & Grant, L. D.) 254 (Karger, Basel, 1975)
18 Sar, M. & Stumpf, W. E. *Cell Tissue Res.* **203**, 1 (1979).
19 Sar, M. & Stumpf, W. E. *J. Histochem. Cytochem.* (in the press)
20 Stumpf, W. E. & Sar, M. *Meth. Enzym.* **36**, 135 (1975)
21 Straus, W. *J. Histochem. Cytochem.* **27**, 1349 (1979)
22. Pickel, V. M., Joh, T. H. & Reis, D. J. *J. Histochem. Cytochem.* **24**, 792 (1976)
23. Stumpf, W. E. & Sar, M. in *Modern Pharmacology–Toxicology* Vol. 8 (ed. Pasqualini, J.) 41 (Dekker, New York, 1976)
24 Heritage, A. S., Grant, L. D. & Stumpf, W. E. *J. comp. Neurol.* **176**, 607 (1977)
25 Hökfelt, J., Fuxe, K., Goldstein, M. & Johansson, O. *Brain Res.* **66**, 235 (1974).
26 Stefano, F. J. E., Donoso, A. O. & Cuner, J. *Acta physiol. latinoam.* **15**, 425 (1965).
27 Donoso, A. O. & Stefano, F. J. E. *Experientia* **23**, 665 (1967).
28 Crowley, W. R., O'Donohue, T. L., Wachtbecht, H. & Jacobowitz, D. M. *Brain Res.* **154**, 345 (1978)

29. Beattie, C. W., Rodgers, C. H. & Soyka, L. F. *Endocrinology* **91,** 276 (1972).
30. Luine, V. N. & McEwen, B. S. *J. Neurochem.* **28,** 1221 (1977).
31. Nicholson, G., Greely, G., Humo, J., Youngblood, W. & Kizer, J. S. *Endocrinology* **103,** 559 (1978).
32. Martinovic, J. & McCann, S. *Endocrinology* **100,** 1206 (1977).
33. Negro-Vilar, A., Ojeda, S. R., Davis, J. P. & McCann, S. *Endocrinology* **105,** 86 (1979).
34. Simpkins, J. W., Huang, H. H., Davis, J. P. & Meites, J. *Biol. Reprod.* **20,** 625 (1979).
35. Kalra, S. P., Kalra, P. S., Chen, C. L. & Clemens, J. A. *Acta endocr.* **89,** 1 (1978).
36. Kun, L. *Rev. Rom. med. Endocr.* **16,** 263 (1978).
37. Day, T. A., Blessing, W. & Willoughby, J. O. *Brain Res.* **193,** 543 (1980).

Enkephalins co-exist with oxytocin and vasopressin in nerve terminals of rat neurohypophysis

R. Martin & K. H. Voigt

Sektion Elektronenmikroskopie und Labor für Neuroendokrinologie,
Universität Ulm, D-7900 ULM, FRG

Evidence is accumulating that opiates inhibit the release of oxytocin and vasopressin by acting on nerve terminals in the neurohypophysis[1,2]. Extracts of neurohypophysis have been shown to contain substantial amounts of Met- and Leu-enkephalin[3,4], and Leu-enkephalin-immunoreactive (IR) nerve fibres originating in the magnocellular hypothalamic nuclei have been described in the neural lobe, where the two hormones are secreted[3]. We have compared the distribution of oxytocin, vasopressin and enkephalin immunoreactivity (IR) in the neurohypophysis of the rat, and report here that Met-enkephalin-IR is invariably associated with nerve terminals that contain oxytocin-IR whereas the terminals that contain vasopressin-IR often, but not invariably, are Leu-enkephalin immunoreactive.

Neural lobes of five normal white rats and four Brattleboro rats homozygous for hypothalamic diabetes insipidus were fixed with paraformaldehyde and glutaraldehyde and embedded in Epon. Resin was removed from serial 0.5-μm sections and Sternberger's unlabelled antibody–enzyme method applied for light microscopic immunocytochemistry[5] (antibodies are listed in Table 1). Serial ultra-thin and semi-thin sections were cut to examine the fine structure of immunoreactive terminals—the semi-thin section was processed for immunocytochemistry and the ultra-thin section stained with uranium and lead salts for electron microscopy[6]. Cross-reactivity of antibodies was examined in preabsorption controls and in radioimmunoassay (RIA) (Table 1, Fig. 1).

Affinity-purified antibodies against oxytocin and vasopressin recognized two populations of nerve terminals[7] (Fig. 2*b*, *c*, *e*). In their fine structure vasopressin terminals differed from oxytocin terminals by having more densely packed neurosecretory granules. Nerve terminals in neural lobes from Brattleboro rats reacted with oxytocin antibody but not vasopressin antibody.

Preabsorption tests showed that, of the three enkephalin antibodies, only affinity-purified Met-enkephalin antibody I specifically recognized Met-enkephalin (Fig. 1*a*, *b*). This was confirmed by RIA with several enkephalin analogues and other peptides (Table 1). Met-enkephalin antibody II cross-reacted with Leu-enkephalin (Fig. 1*c*, *d*) and Leu-enkephalin antibody cross-reacted with Met-enkephalin. This specificity was not affected by addition of Leu-enkephalin or Met-enkephalin in decreasing concentrations or by varying the dilution of the antiserum. Thus, presumably, nerves that were stained by all three antibodies contained Met-enkephalin-like material, and nerves stained by the two cross-reacting antibodies only contained Leu-enkephalin-like material. However, it cannot be ruled out that nerves exhibiting Met-enkephalin-IR contained Leu-enkephalin-like material. Furthermore, it is possible that all the antibodies cross-reacted with larger peptides containing the sequence against which the antibodies were raised—in preabsorption controls and RIA the Met/Leu-enkephalin antibodies but not the Met-enkephalin antibody I recognized the larger dynorphin[9].

When enkephalin antibody I, oxytocin and vasopressin antibodies were applied in turn to serial sections, Met-enkephalin-IR nerves were indistinguishable in size, shape and position from most of the oxytocin-IR terminals in a subsequent section (Fig. 2*a*, *c*, *f*; Fig. 3), but not from vasopressin-IR terminals (Fig. 2*a*, *b*, *g*). Fine structure revealed that the immunoreactive structures were nerve terminals or varicosities filled with electron-dense granules >100 nm in diameter; oxytocin-IR terminals were indistinguishable from Met-enkephalin I-IR terminals. Among terminals with diameters >1 μm in the neurohypophysis of a Brattleboro rat (Fig. 3), 111 out of 117 Met-enkephalin I-IR terminals coincided with oxytocin-IR terminals in consecutive sections. Six enkephalin-IR terminals without corresponding oxytocin-IR structures were all small and may not have appeared in the next section.

Oxytocin-IR terminals were recognized by Met-enkephalin antibody II (Fig. 2*d*, *c*, *h*) and by the Leu-enkephalin antibody which cross-reacted with Met-enkephalin. We therefore assume that oxytocin terminals contain Met-enkephalin-like material. As we had no specific Leu-enkephalin antibody we cannot exclude the possibility that they also contain Leu-enkephalin-like material. In well stained preparations there was nearly one to one correspondence between Met-enkephalin-IR and oxytocin-IR in consecutive sections.

Met-enkephalin antibody II, and the Leu-enkephalin antibody, which cross-reacted with both Leu- and Met-enkephalin respectively, stained certain vasopressin-IR terminals in addition to oxytocin-IR structures (Fig. 2*b*–*d*, *h*, *i*). The Leu/Met-enkephalin antibodies probably recognized Leu-enkephalin-like material but not Met-enkephalin-like material in vasopressin terminals, for the specific Met-enkephalin antibody I did not stain the corresponding structures. However, correspondence

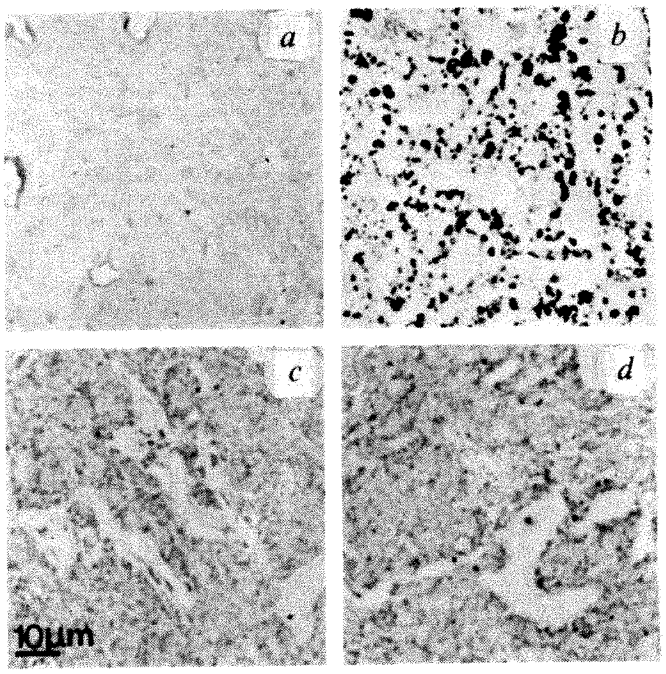


Fig. 1 Preabsorption tests using neural lobe sections with the two Met-enkephalin antibodies. *a*, *b*, Antibody I; *c*, *d*, antibody II. *a*, *c*, Preabsorption by Met-enkephalin; *b*, *d*, by Leu-enkephalin. When preabsorbed by Met-enkephalin immunoreaction was absent, that is, both antibodies had recognized the antigen and binding sites were blocked. Antibody I, in contrast to antibody II, had not recognized Leu-enkephalin and was reacting with the tissue (*b*). ×580.

0028–0836/81/060502—03$01.00

Table 1 Antibodies and specificity tests

	Antibodies				
	Anti-Met-enk. I (R26)*	Anti-Met-enk. II†	Anti-Leu-enk.†	Anti-oxytocin (02C)	Anti-vasopressin (125)
Working dilution	1:2,000	1:4,000	1:4,000	1:3,200	1:1,600
Purification by affinity chromatography	+	–	–	+	+
Recognized antigens	Met-enk.	Met-enk. Leu-enk. Dynorphin	Leu-enk. Met-enk. Dynorphin	Oxytocin	Vasopressin α-MSH
Antigens not recognized	Leu-enk. Oxytocin Dynorphin	Oxytocin Arg-vasopressin Lys-vasopressin Vasotocin	Oxytocin Arg-vasopressin	Arg-vasopressin Met-enk. Leu-enk. Neurophysin I/II	Oxytocin Met-enk. Leu-enk. Neurophysin I/II

Antibodies were raised against synthetic peptides. Cross-reactivities with the listed antigens were examined in preabsorption controls by incubating the antibodies for 24 h with the peptides (20 µl of the antibody solution at half the working dilution +5 µg peptide in 5 µl phosphate-buffered saline (PBS) + 15 µl PBS). Absence of staining in neural lobe sections after application of the antibody–antigen mixtures indicated that the added peptide had been recognized by the antibody (Fig. 1).

* In a RIA[8] (using ^{3}H-Met-enkephalin as tracer and unpurified antiserum R 26, 1:16,000), the Met-enkephalin antibody I did not cross-react with the amino acids methionine and leucine, nor the sequences Phe-Met, Phe-NLeu, Phe-Leu, Gly-Gly-Phe-Met, Gly-Gly-Phe-NLeu, and the peptides α-melanocyte-stimulating hormone (α-MSH), β-endorphin, lipotropin hormone (β-LPH), ACTH_{1-24}, oxytocin and vasopressin. It cross-reacted less than 0.5% with dynorphin or Leu-enkephalin, but 11% with NLeu-enkephalin.

† It was not possible to isolate a specific population of these antibodies by affinity chromatography. The purification procedures and cross-reactivities of oxytocin and vasopressin antibody have been described elsewhere[7]. The vasopressin antibody stained intermediate lobe cells which indicated cross-reactivity with α-MSH[7]. When preincubated with synthetic α-MSH, staining of intermediate lobe was abolished but staining of neural lobe terminals was not impaired. Application of two α-MSH antibodies[6] did not reveal α-MSH-IR in the neural lobe.

between Leu-enkephalin-IR and vasopressin-IR was less complete than between Met-enkephalin-IR and oxytocin-IR. In many regions vasopressin terminals were not recognized by Met/Leu-enkephalin antibodies. The correlation was better in larger than in smaller terminals (see Fig. 2). In Brattleboro rats in which only oxytocin-IR terminals were found, all these terminals seemed to react with all three enkephalin antibodies. No additional stained structures were observed when the immunoreactivity of the two Met/Leu-enkephalin antibodies was compared with that of the Met-enkephalin antibody I.

Immunoreactive terminals used for comparison had diameters of 1–12 µm, so that most of them would appear in consecutive 0.5-µm sections. In one series, 91% of the enkephalin-IR terminals (Met-enkephalin antibody II) coincided with either oxytocin or vasopressin terminals; the remaining 9% were all rather small and were presumably cut tangentially. Thus there was no indication of separate enkephalinergic nerves.

Our observation of nerves with coexisting Met(Leu)-enkephalin/oxytocin-IR and Leu-enkephalin/vasopressin-IR does not necessarily contradict the report of enkephalin-IR nerves in the rat neurohypophysis[3], for that study did not include differential staining with enkephalin, oxytocin and vasopressin antibodies. However, the cell bodies of those enkephalin nerves were located in hypothalamic nuclei that contain oxytocin and vasopressin neurones. In dehydrated rats the decrease in vasopressin and Leu-enkephalin content[3,10], but not of thyrotropin-

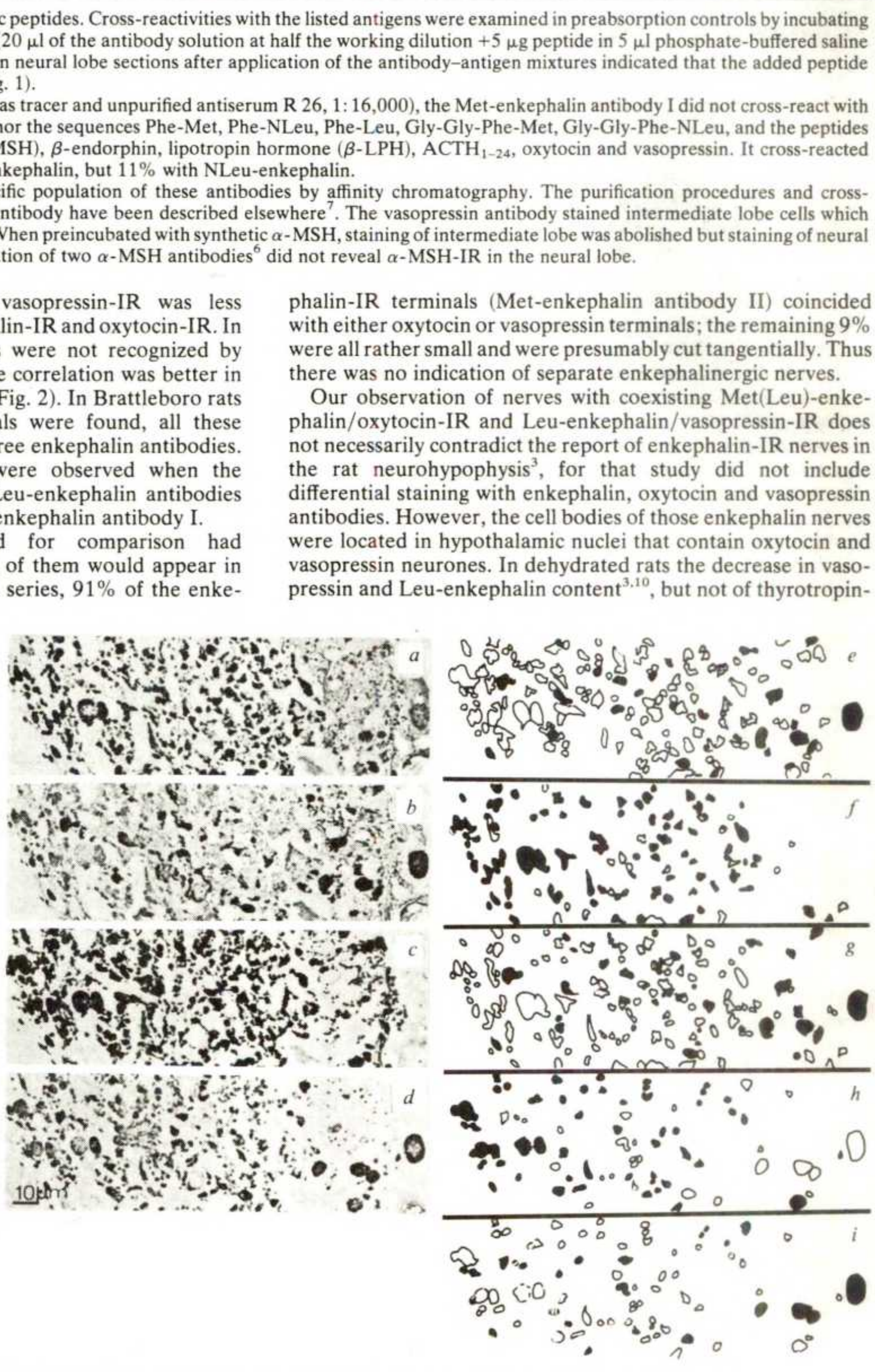


Fig. 2 Four consecutive 0.5 µm-thick sections through a neural lobe showing: *a*, Met-enkephalin I-IR; *b*, vasopressin-IR; *c*, oxytocin-IR; *d*, Met-enkephalin II-IR. ×350. *e–i*, The immunoreactive structures on two of the sections were compared: *e*, vasopressin-IR (*b*) (black spots) compared with oxytocin-IR (*c*) (contours). The vasopressin antibody recognized other terminals than the oxytocin antibody. *f*, Met-enkephalin I-IR (*a*) (contours) compared with oxytocin-IR (*c*). Coincidence of immunoreactive structures is marked by black filling. The Met-enkephalin antibody I recognized most oxytocin terminals. *g*, Met-enkephalin I-IR (*a*) (contours) compared with vasopressin-IR (*b*) (black spots). The Met-enkephalin antibody I did not recognize vasopressin terminals. *h*, Met-enkephalin II-IR (*d*) (contours) compared with oxytocin-IR (*c*); coincidence marked by black filling. The Met-enkephalin antibody II recognized oxytocin terminals. *i*, Met-enkephalin II-IR (*d*) (contours) compared with vasopressin-IR (*b*); coincidence marked by black filling. The Met-enkephalin antibody II also recognized vasopressin terminals.

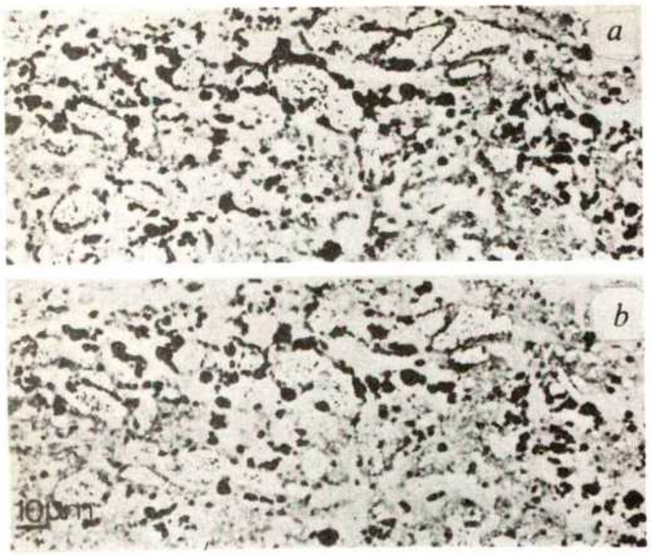


Fig. 3 Two serial sections through a Brattleboro rat neural lobe showing *a*, oxytocin-IR, and *b*, Met-enkephalin I-IR. Most immunoreactive terminals coincide on the two sections. ×400.

releasing hormone or of somatostatin stores[3], agrees with coexistence of the two peptides in the same nerves. Also, in homozygous Brattleboro rats enkephalin content was distinctly lower than that in heterozygote littermate controls[3,10]. Thus it is possible that in this vasopressin-deficient strain a vasopressin-related enkephalin system is impaired.

Our study suggests that enkephalins are co-peptides of the oxytocin and vasopressin neurosecretory systems and that Met(Leu)- and Leu-enkephalin occur in separate neurones. Whether the two peptides are stored in the same secretory granule is not known. The secreted enkephalins may act on autoreceptors on their own nerve fibre or may influence neighbouring terminals, interfering with oxytocin and vasopressin release.

Differential distribution of Met- and Leu-enkephalin-IR has been found in another neuronal system[11]. However, the existence of separate neurones for the two enkephalins has been questioned by the characterization of a putative common precursor of Met-enkephalin and Leu-enkephalin in the adrenal medulla[12]. However, so far only peptides containing either the Met- or Leu-enkephalin sequence have been extracted from brain tissue[13,14] or pituitaries[9,15].

We thank Dr F. W. van Leeuwen for purified oxytocin and vasopressin antibodies. Anti-Met-enkephalin II and anti-Leu-enkephalin were from the Immuno Nuclear Corporation. Anti-Met-enkephalin I was from K.H.V. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 87, B2). K.H.V. is a recipient of a Heisenberg Stipendium.

Received 21 July; accepted 5 November 1980.

1. Clarke, G., Wood, P., Merrick, L. & Lincoln, D. W. *Nature* **282,** 746–748 (1979).
2. Iversen, L. L., Iversen, S. D. & Bloom, F. E. *Nature* **284,** 350–351 (1980).
3. Rossier, J. *et al. Nature* **277,** 653–655 (1979).
4. Kumar, M. S. A., Chen, C. L. & Muther, T. F. *Life Sci.* **25,** 1687–1696 (1979).
5. Weber, E., Voigt, K. H. & Martin, R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 6134–6138 (1978).
6. Martin, R., Frösch, D. & Voigt, K. H. *Gen. comp. Endocr.* **42,** 235–243 (1980).
7. Van Leeuwen, F. W., de Raay, C., Swaab, D. F. & Fisser, B. *Cell Tissue Res.* **202,** 189–210 (1979).
8. Voigt, K. H., Weber, E. & Martin, R. in *Structure and Activity of Natural Peptides* (eds Voelter, W. & Weitzel, G.) (de Gruyter, Berlin, 1980).
9. Goldstein, A., Tachibana, S., Lowney, L. J., Hunkapillar, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6666–6670 (1979).
10. Mata, M. M., Gainer, H. & Klee, W. A. *Life Sci.* **21,** 1159–1162 (1977).
11. Larsson, L. J., Childers, S. & Snyder, S. H. *Nature* **282,** 407–410 (1979).
12. Lewis, R. V. *et al. Science* **208,** 1459–1461 (1980).
13. Stern, A. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 6680–6683 (1979).
14. Huang, W. Y., Chang, R. C. C., Kastin, A. J., Coy, D. H. & Schally, A. V. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6177–6180 (1979).
15. Kangawa, K. & Matsuo, H. *Biochem. biophys. Res. Commun.* **86,** 153–160 (1979).

Milk ejection in a marsupial, *Macropus agilis*

Dennis W. Lincoln & Marilyn B. Renfree

School of Environmental and Life Sciences, Murdoch University, Western Australia 6150

Marsupials give birth to young which are almost embryonic in form, and for the first few months of postnatal life they remain continuously attached to the nipples of the lactating mammary glands. The immature appearance of the neonate and the presence of an elaborate musculature in the wall of the pouch led observers to suggest that the contraction of the pouch functioned to compress the lactating mammary glands and thus force milk into the mouth of the young[1]. This view was dismissed when Enders[2] demonstrated that milk was not expressed in the opossum, *Didelphis virginiana*, following electrical stimulation of the motor nerves supplying the muscles of the pouch, and he implied that milk was obtained solely by the sucking of the young. In most eutherian mammals very little milk can be extracted unless facilitated by an oxytocin-induced contractionof the myoepithelial cells surrounding the alveoli of the mammary gland. It is not known whether such a process of milk ejection exists in a marsupial, although oxytocin has been identified in the posterior pituitary gland[3,4] and has been used as an aid in the collection of milk[5]. We recorded intramammary pressure from the lactating mammary gland of the agile wallaby, *Macropus agilis*, and examined the pressure responses evoked by injection of oxytocin and electrical stimulation of the presumptive oxytocin neurones. We report here three striking differences from those reported for eutherian mammals: the mammary gland of the wallaby, particularly during the early part of lactation (<100 days), was more sensitive to oxytocin, produced higher intraductal pressures and contracted at lower frequencies of brain stimulation (4 pulses per s). We may therefore conclude that the wallaby, far from discarding during its evolution the milk-ejection response to oxytocin, has refined the response to permit it to feed simultaneously both newly born pouch young (<1 g) and young at foot (>2,500 g).

Studies were made of 18 adult female agile wallabies maintained in the Native Fauna Research Unit of Murdoch University, Western Australia. The agile wallaby is a continuous and opportunistic breeder found abundantly in the wild in the northern parts of Australia and New Guinea[6,7]. Pregnancy lasts for less than 1 month and the neonate (~650 mg) attaches to one of the four nipples in the pouch and induces a lactation confined to that mammary gland for about 350 d. Lactation is characterized by three stage: (1) pouch young continuously attached to the nipple (0–100 days); (2) pouch young detached from the nipple (80–210 days), and (3) young at foot and sucking occasionally (190–360 days). Evacuation of the pouch occurs at ~200 days—this triggers reactivation in the growth of the embryo in the uterus with the birth of the baby 26 days later[7]. The neonate attaches to one of the three unused nipples and as a result a second lactation develops in parallel to the previous one, but lagging by over 200 days.

After capture, the pouch young were removed and aged[6], and the mothers (10–15 kg) were anaesthetized with intravenous sodium pentobarbitone (30 mg per kg). Intramammary pressures were recorded from one or more of the 20–30 galactophores in the nipples of the lactating mammary glands. Recordings were made from glands, which according to the age of the dependent young, varied over 38–350 days of lactation.

Injections of oxytocin, either into a catheterized tail vein or external jugular vein, promoted a rise in intramammary pressure after 23–60 s, depending on the dose of oxytocin and the

0028-0836/81/060504—03$01.00

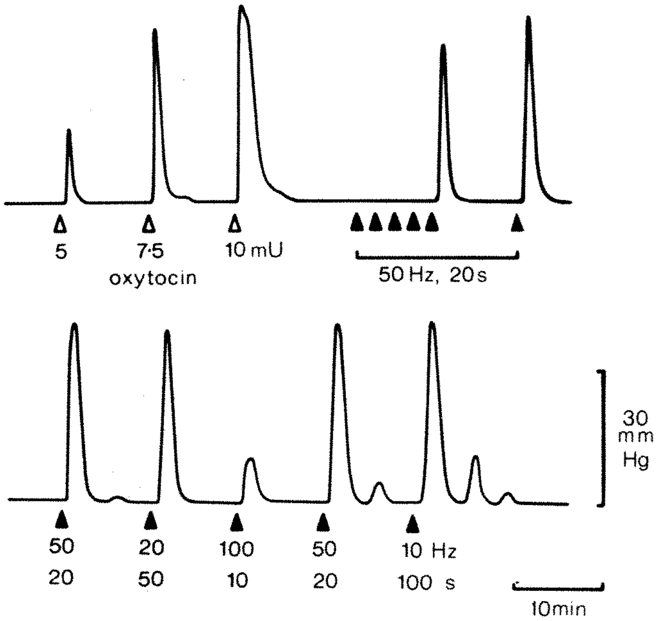


Fig. 1 Recordings of pressure from a galactophore of a mammary gland of an agile wallaby *Macropus agilis* at day 82 of lactation, illustrating the response to injections of oxytocin and electrical stimulation of presumptive oxytocinergic neurones in the region of the basal hypothalamus. The wallaby was anaesthetized with sodium pentobarbitone (30 mg per kg) and the galactophore, one of 27 in the nipple, was cannulated with a 25-gauge needle and connected to a pressure transducer filled with 1% sodium citrate. The two traces are continuous. In the upper trace the first three peaks are the responses to three bolus injections of oxytocin into the external jugular vein. A stimulating electrode was then lowered into the basal part of the hypothalamus from a point of entry in the skull 28 mm anterior to the bregma suture. A 50 Hz × 20 s pulse train (1 ms × 700 μA biphasic pulses) (solid triangles) was applied at intervals as the electrode was lowered, until a position was obtained within the brain at which milk-ejection was induced. The lower trace illustrates a stimulation protocol in which 1,000 pulses were applied at frequencies ranging from 10 to 100 Hz, at 10-min intervals.

lactational age of the mammary gland. Peak pressures occurred after 60–90 s and baseline pressures returned after 120–180 s. The threshold dose of oxytocin for the production of a mammary contraction was 2–15 mU, and was substantially lower in recordings from mammary glands in early lactation (< 100 d). The dose–response curve was steep and the pressure responses attained a plateau of 30–60 mm Hg after injections of ≥20–30 mU oxytocin (Figs 1, 2).

Electrical stimulation of the hypothalamo-neurohypophysial tract in the region of the basal hypothalamus with a 50 Hz × 20 s pulse train was most effective at promoting milk ejection. The pressure wave evoked by stimulation was slightly more progressive in onset than that produced by a bolus injection of oxytocin, although the peak of the pressure wave was slightly more abrupt. A 50-Hz frequency is optional for the induction of oxytocin release in the rabbit and rat and insufficient oxytocin is released in these species to promote milk ejection at ≤20 Hz (refs 8, 9). Surprisingly, when 1,000 pulses were applied to the wallaby in the form of a 10 Hz × 100s pulse train the milk-ejection response was equal in amplitude to that seen with higher frequencies of stimulation. Subsequently, very effective increases in pressure were obtained with prolonged stimulation at 4.5 Hz and some evidence of oxytocin release was seen with even lower frequencies (Fig. 2).

Marsupials have evolved a highly successful reproductive strategy when compared with eutherian mammals. A key feature of their success seems to be the refinement of the milk-ejection response to oxytocin to permit the mother to nourish the very immature young continuously in the months immediately after birth and, in some species, to allow her intermittently to feed an older sibling on a second lactating mammary gland. Two observations make this a practical possibility: release of oxytocin in response to low-frequency electrical stimulation and the extreme oxytocin sensitivity of the mammary glands in early lactation. In eutherian mammals, the posterior pituitary gland contains a calcium-dependent process of spike facilitation where action potentials entering the nerve terminals in the gland, at ≥30 Hz, release 100–1,000 times more oxytocin than impulses arriving at ≤10 Hz, and such facilitation is used to generate a pulsatile pattern of hormone release[9]. Frequency facilitation was not observed in the wallaby, which suggests that the process of hormone release differs in some fundamental respect from that in the rat. However, the observation of milk ejection in response to electrical stimulation at 5 Hz raises the possibility that low levels of electrical activity in the oxytocinergic neurones during the early stages of lactation such as might be evoked by the activities of a small pouch young, could result in a small but sustained release of oxytocin sufficient to contract the highly sensitive mammary gland. Such a release would serve in the absence of a tight teat sphincter to produce 'continuous' milk-ejection, with mammary contractions recurring at intervals of ~4 min (Figs 1, 2). Furthermore, the pressures generated in these young glands at milk ejection are fivefold higher than those recorded in a rat and are sufficient to cause milk to leak from the nipple in the absence of an attached young. Thus, the original idea that milk was forced into the neonate may not be so far from the truth, although the mechanism is one of milk ejection rather than contraction of the pouch muscles. The decline in oxytocin sensitivity as lactation progresses makes it possible to run two lactations in parallel, but >200 days out of step. A low level of oxytocin, as discussed above, would not contract a mammary gland feeding a young at foot. Clearly, however, a juvenile of >3,000 g sucking on the enlarged nipple of a mammary gland in late lactation must generate an enormous sensory stimulus compared with that produced by a neonate of <1 g. In such circumstances it seems plausible that a large release of oxytocin could occur and cause a contraction of the older mammary gland. One consequence, though of little

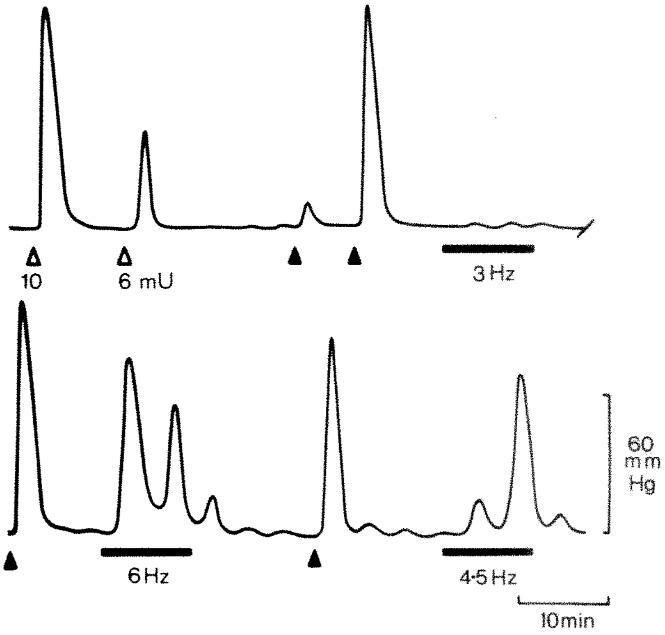


Fig. 2 Intramammary pressure recordings from an agile wallaby at day 116 of lactation. The recording technique and the presentation of the data are as described in Fig. 1 legend. The two traces are continuous, and illustrate the relative effectiveness of prolonged low-frequency stimulation at 3–6 Hz (solid bar) compared with a short burst of high-frequency stimulation (50 Hz × 20 s) (solid triangles) on the induction of milk ejection. Note that prolonged stimulation produced multiple contractions with pressure peaks at intervals of ~4 min.

significance, as the mammary pressure response peaks at 30–60 mm Hg with oxytocin injections >30 mU, is that the older young may give the younger sibling an additional meal.

We acknowledge the financial support of the Australian Research Grants Committee (D.I. 7515759), and the Royal Society for a Commonwealth Bursary.

Received 5 November; accepted 18 November 1980.

1. Morgan, J. *Trans. Linn. Soc. Lond.* **16,** 61–84 (1833).
2. Enders, R. K. *Symp. zool. Soc. Lond.* **15,** 195–203 (1966).
3. Ferguson, D. R. & Heller, H., *J. Physiol., Lond.* **180,** 846–863 (1965).
4. Chauvet, M. T., Hurpet, D., Chauvet, J. & Acher, R. *Nature* **287,** 640–642 (1980).
5. Gross, R. & Bollinger, A. *Am. J. Dis. Childh.* **98,** 102–109 (1959).
6. Kirkpatrick, T. H. & Johnson, P. M. *Qd J Agric. Animal Sci.* **26,** 691–698 (1969).
7. Merchant, J. C. *Aust. Wildl. Res.* **3,** 93–103 (1976).
8. Harris, G. W., Manabe, Y. & Ruf, K. B. *J. Physiol., Lond.* **203,** 67–81 (1969).
9. Lincoln, D. W., Clarke, G., Mason, G. A. & Dreifuss, J. J. in *Neurohypophysis* (eds Moses A. M. & Share, L.) 101–112 (S. Karger, Basal, 1977).

A neuropterous larva uses an allomone to attack termites

J. B. Johnson & K. S. Hagen

Division of Biological Control, Department of Entomological Sciences, University of California, Berkeley, California 94720

Larvae of *Lomamyia* spp. (Neuroptera: Berothidae) have been tentatively associated with several types of prey since they were first identified[1]. In all cases, the assumed prey lived in a confined habitat—an ant nest or termite gallery[2]. *Lomamyia latipennis* Carpenter has been the most thoroughly studied of these unusual Neuroptera. The immature states were described[3,4] and the association with termites strengthened using *Zootermopsis angusticollis* Hagen (Isoptera: Kalotermitidae) as prey for a group of larvae[4]. A low survival rate of the larvae suggested that another termite might be the normal host. We now report field and laboratory findings which demonstrate the predator–prey relationship between *L. latipennis* and the subterranean termite *Reticulitermes hesperus* Banks (Isoptera: Rhinotermitidae) and describe their unique method of attacking prey with an aggressive allomone.

During the summers of 1979 and 1980 field studies and collections were conducted in Gates Canyon, Solano County, California. Three significant field observations were made concerning the association of *L. latipennis* and *R. hesperus.* On 18 June 1979, a cluster of eight hatched berothid eggs was found hanging on a single stalk from a projection of sapwood on an oak log (*Quercus* sp.) infested with *R. hesperus.* If the larvae behaved like those observed in the laboratory, they crawled up the stalk and across the log surface entering holes and crevices, seeking the termites in the log. When the log was chopped open a well fed first instar berothid larva was found. Then, on 14 June 1980 an adult female *L. latipennis* emerged from a stump of elderberry (*Sambucus coerulea* Raf.) which had been brought into the laboratory about 1 month earlier as a source of *R. hesperus.* No evidence of berothids was found in or near colonies of *Z. angusticollis* or *Zootermopsis nevadensis* Hagen, the only other species of termites recorded in northern California.

Nine female *L. latipennis* were collected under mixed UV/visible light on 2 June 1979 and 9 June 1980. They deposited fertile eggs in the laboratory when fed artificial honeydew consisting of dried yeast (*Kluyveromyces fragilis*), autolysed brewers yeast (*Saccharomyces cerevisiae*) and honey (6 g : 1 g : 10 g). After hatching, individual larvae were placed in 3-cm^3 glass vials half-filled with frass from the termite colonies. Live termites were supplied daily from source colonies maintained in the laboratory. All three species of termite were tested as prey for *L. latipennis* larvae. Only larvae fed *R. hesperus* successfully completed their larval development.

Observations of *L. latipennis* larvae in laboratory colonies of *R. hesperus* indicated that they are true termitophiles, living and moving freely among the termites in their galleries. In the colonies and rearing vials the larvae generally exhibited some caution when near the much more heavily sclerotized termites, but no larva was observed to be attacked. This pattern of coexistence changed only when the larvae were ready to feed.

First instar *L. latipennis* larvae ($\sim$0.07 mg), which are much smaller than *R. hesperus* workers ($\sim$2.5 mg), displayed a consistent attacking behaviour. A larva repeatedly approached and retreated until the tip of its abdomen was directed at the termite's head. The apex of the abdomen was then lifted and waved past the termite's face, apparently without contact. The termite apparently was not repelled, as it made no obvious effort to escape. One to three minutes later it was incapacitated, lying supine, with its legs moving irregularly. Two to five minutes after the collapse, motion ceased. During this time, the berothid larva continued to advance and withdraw. Feeding occurred only after the prey had become still. If not fed upon, the termite's heart would continue to beat for long as 3 h, showing that the initial immobility was merely paralysis, though death ultimately occurred. The termite may also be preserved by the allomone, but initial tests were inconclusive.

Third instar *L. latipennis* larvae behaved similarly, but the obvious wave of the abdomen was not seen and individuals were observed to kill as many as six *R. hesperus* simultaneously, in the rearing vials. Apparently, the larger larvae produce enough of the allomone to eliminate the need for such precise application to the prey. Second instar larvae are sessile and do not feed[4]. As with the active instars, they were not attacked by the termites, although they were exposed and apparently defenceless.

Soldiers and reproductives of *R. hesperus* were affected in the same manner as the workers. Individual 'pseudoworkers' of the *Zootermopsis* spp. were not affected in this way by *L. latipennis* larvae, although the larvae could feed, with apparent difficulty, on freshly killed *Zootermopsis.* This difference could be due to physiological differences between the termites or simply a matter of dosage, because the *Zootermopsis* spp. are much larger ($\sim$8 to >50 mg). If it is a matter of dosage this could explain the partial success of the group rearing using *Z. angusticollis* as prey[4].

Two compartment test chambers were used to verify that contact between the larva and termite was not required. In each case, a berothid larva and one *R. hesperus* were put into one compartment and a termite into the other. The first type of test chamber was a 0.7-cm^3 glass vial divided longitudinally by a strip of filter paper. In 10 tests the second termite was invariably incapacitated within a few minutes after the first had been attacked by the berothid. The second variety of test chamber was a 0.5-cm^3 glass tube divided transversely by a disk of cork pierced by a piece of organdy screened capillary tubing. When the prey termite had been attacked, a syringe was used to pull a volume of air, equivalent to that in the first chamber, into the second, as a capillary tube vent permitted air to enter the first. In 7 out of 20 trials, the test termite was also affected. The inconsistent results were attributed to the relatively crude construction of the test chamber and variability in the timing of the suction. In 20 tests without a berothid larva, none of the 40 termites exhibited any signs of distress.

As it can pass through a filter paper barrier and be moved by an air current, the toxin must be active as a gas or an aerosol. This observation of an allomone used in predation is unique among the insects. The only arthropod previously known to use an allomone in this way seems to be the mite, *Womersia strandtmanni* Wharton (Acarina: Leeuwenhoekiidae)[5]. Because the *L. latipennis* allomone apparently lacks a defensive function, we believe that its primary purpose is aggressive, and we propose the term aggressive allomone to describe this situation. Theoretically, it is possible, and in fact seems probable, that the berothid's allomone evolved from a defensive allomone which was necessary early in their association with *R. hesperus*, before their termitophilia evolved to its current state.

0028–0836/81/060506—02$01.00

While rearing *L. latipennis* larvae, we noted that Psocoptera (Psocidae) lived in the vials unaffected by the allomone. To test the specificity of the toxin further small Diptera (Drosophilidae) and Hymenoptera (*Coccophagus cowperi* Girault, Aphelinidae and *Metaphycus funicularis* Annecke and *Metaphycus stramineus* Compere, Encyrtidae) were placed in the test chamber. None of these insects, or other berothid larvae, was affected by the allomone.

Motivated by the unique nature of the allomone, its specificity and the importance of the target insects, efforts are now directed at establishing laboratory cultures of *L. latipennis* and isolating and identifying the active compound(s). Synthesis of the allomone would permit detailed study of its activity and specificity and could yield a valuable new means of controlling termites.

We thank Dr Peter Duelli for his assistance and discussions. This research was supported, in part, by the USDA Western Regional Biological Control Project (W-84).

Received 11 August; accepted 28 November 1980.

1. Smith, R. C. *Ann. ent. Soc. Am.* **16,** 129–151 (1923).
2. Gurney, A. B. *Psyche* **54,** 145–169 (1947).
3. Tosch, C. A. *Pan-Pacific Ent.* **40,** 21–26 (1964).
4. Tauber, C. A. & Tauber, M. J. *Can. Ent.* **100,** 623–629 (1968).
5. Huber, I. *Acarologia* **20,** 112–115 (1979).

Circulating angiotensin II and adrenal receptors after nephrectomy

G. Aguilera, A. Schirar, A. Baukal & K. J. Catt

Endocrinology and Reproduction Research Branch,
National Institute of Child Health and Human Development,
National Institutes of Health, Bethesda, Maryland 20205

Mineralocorticoid secretion is predominantly controlled by the octapeptide angiotensin II, which exerts trophic actions on the adrenal glomerulosa and acute regulatory effects on aldosterone biosynthesis. The trophic actions include stimulation of angiotensin II receptors and enzymes of the aldosterone biosynthetic pathway[1,2], with corresponding enhancement of the aldosterone secretory capacity of the adrenal gland. The positive regulatory action of angiotensin II on its adrenal receptors occurs with elevations of the circulating peptide concentration within the physiological range and probably contributes to the increased sensitivity of the adrenal during sodium deficiency[3,4]. In this action, angiotensin II differs from other hormones which decrease their target-cell receptors[5–7]. However, the increase in adrenal angiotensin II receptors following nephrectomy has been interpreted as evidence for a tonic down-regulating effect of angiotensin II on its adrenal receptors[8]. To clarify these conflicting views we evaluated the effects of nephrectomy on adrenal angiotensin II receptors in relation to blood angiotensin II and plasma electrolyte levels. We show here that hyperkalaemia contributes markedly to the post-nephrectomy increase in adrenal angiotensin II receptors, and that circulating angiotensin II levels persist for an unexpectedly long period after nephrectomy, presumably due to tissue generation of the octapeptide.

Adult male Sprague–Dawley rats were bilaterally nephrectomized and groups killed at intervals up to 48 h. Angiotensin II receptors were measured in the zona glomerulosa by binding assay of adrenal capsule homogenate as described previously[9] and blood samples were obtained for analysis of plasma renin activity (PRA)[10], blood angiotensin II (ref. 11) and serum

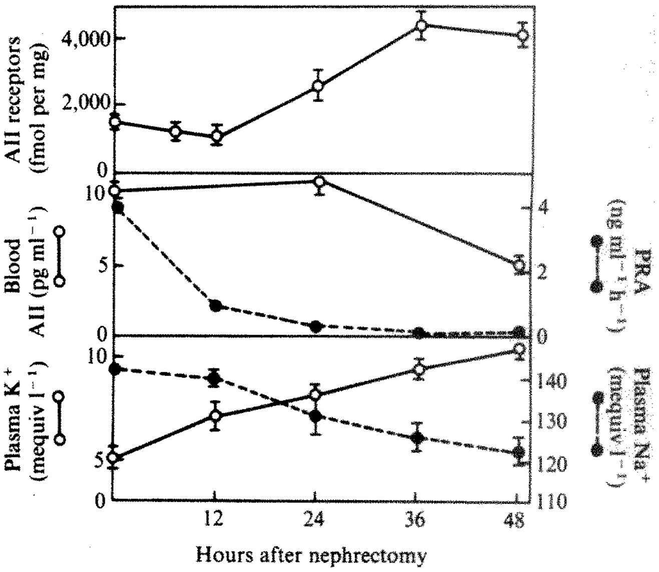


Fig. 1 Effects of nephrectomy on adrenal glomerulosa angiotensin II (A II) receptors (upper panel), plasma renin activity (PRA) and angiotensin II (middle panel) and plasma sodium and potassium (lower panel). The number of angiotensin II receptors was measured in the 300,000*g* particulate fraction of adrenal capsular tissue[9], and data points represent the mean and s.e. of three experiments. The data for PRA and blood angiotensin II are the mean and s.e. of assays performed on 6–15 individual animals. Angiotensin II was assayed in 5–7-ml aliquots of blood, collected in acetone and processed as previously described[11]. The extracts were resuspended in 1 ml of HEPES buffer and angiotensin II concentration was determined by radioimmunoassay (RIA). The ED_{50} of the RIA was 30–40 pg angiotensin II and the lower limit of sensitivity was less than 5 pg, allowing the detection of values as low as 2 pg ml^{-1} of blood.

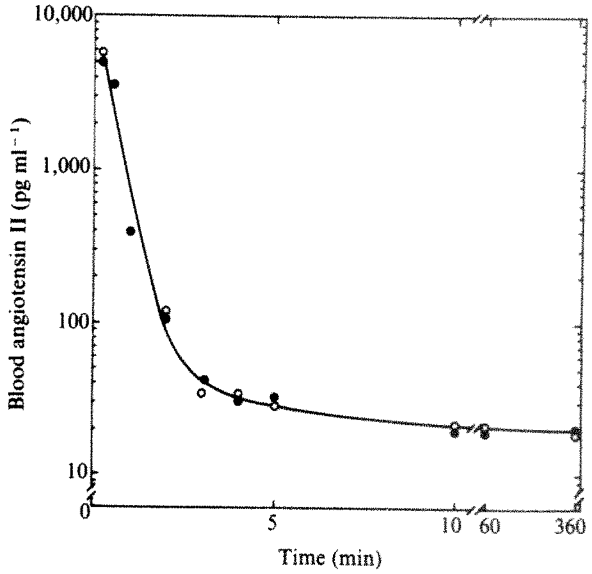


Fig. 2 Rate of disappearance of angiotensin II from the circulation of intact (○) and nephrectomized (●) rats. After injection of 500 ng angiotensin II as a single intravenous injection under ether anaesthesia, animals were killed at timed intervals by decapitation and blood was collected into acetone for extraction and RIA of angiotensin II. Each point represents the mean of values derived from three or four rats.

electrolytes. After nephrectomy, the number of adrenal angiotensin II receptors was unchanged for 12 h, and then increased progressively over the next 24 h to almost double the control level (Fig. 1). As expected, PRA was markedly decreased after nephrectomy and became almost undetectable by 36 h. However, when PRA was reduced by 90% at 24 h, blood angiotensin II levels did not show a commensurate fall, but remained at the levels found in intact rats. Thus, the concentration of blood angiotensin II in 21 nephrectomized rats was 14.7 ± 6.9 pg ml^{-1} (mean ±s.d.) and in 22 normal animals 13.2 ± 6.0 pg ml^{-1}. Furthermore, after 48 h, when PRA was undetectable, there was only a 50% fall in blood angiotensin II concentration. In contrast, low blood angiotensin II levels were found in sodium-loaded rats (1.3 ± 1.5 pg ml^{-1}, $n = 11$) and in animals treated with the converting-enzyme inhibitor, captopril (2.1 ± 2.1 pg ml^{-1}, $n = 9$). In nephrectomized rats, plasma sodium decreased, and plasma potassium increased markedly to levels over 10 mequiv l^{-1} within 48 h (Fig. 1). Elevation of the extracellular potassium concentration is known to exert a direct trophic action on the zona glomerulosa, causing marked increases in the number of angiotensin II receptors and aldosterone secretion[12]. The rise in plasma potassium following nephrectomy would therefore contribute to the observed increase in angiotensin II receptors in the adrenal gland. These results also demonstrate that the increase in angiotensin II receptors observed after nephrectomy cannot be attributed to the removal of a negative effect of circulating angiotensin II on its adrenal receptors.

The unexpected persistence of normal circulating levels of angiotensin II after nephrectomy was not due to a decrease in the metabolic clearance rate of the peptide in the absence of the kidneys. As shown in Fig. 2, the disappearance of angiotensin II from the circulation after a single injection of the peptide was identical in intact and nephrectomized rats. Computer analysis of the data[13] gave half times of 0.46 ± 0.1 and 4.0 ± 0.2 min for the fast and slow components of the disappearance curves. The rapid clearance of angiotensin II from the circulation indicates that angiotensin formation from non-renal renin or renin-like enzymes must be responsible for the circulating peptide observed in anephric animals.

Such extravascular generation of angiotensin II is also suggested by the finding that several tissue extracts contain immunoreactive angiotensin-like material. The tissue content of angiotensin II in adrenal capsule, decapsulated adrenal, aorta and uterus was: 22.2 ± 5, 7.3 ± 2, 11.5 ± 3 and 14.2 ± 2 pg per mg, respectively (mean ±s.e. of 5–10 determinations). In diaphragm extracts, only negligible amounts of the peptide were found. The relatively high content of immunoreactive angiotensin II in the adrenal capsule was not decreased after nephrectomy. The adrenal capsular content of angiotensin II is not solely attributable to receptor-bound peptide, as adrenal angiotensin II was decreased rather than increased after infusion of the peptide (Fig. 3). The absence of increased capsular angiotensin II is not simply the result of receptor down-regulation by the infused peptide[8], because high doses of angiotensin II reduce but do not abolish its adrenal receptors[4].

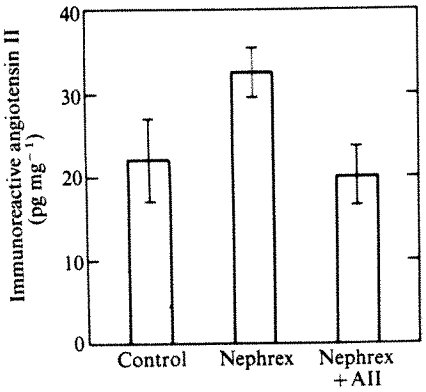


Fig. 3 Immunoreactive angiotensin II in extracts of rat adrenal capsule: effects of nephrectomy and angiotensin II infusion. Angiotensin II was infused at the rate of 200 ng min^{-1} from osmotic minipumps (Alza) implanted intraperitoneally immediately after nephrectomy. Adrenal capsules were rapidly frozen on dry ice, homogenized in 0.01 M acetic acid, boiled for 10 min and extracted in 85% acetone. The extracts were dried down and redissolved in phosphate buffer, and angiotensin II concentration was measured by RIA. Bars represent the mean and s.e. of values obtained for the paired adrenal capsules of 6–10 individual rats.

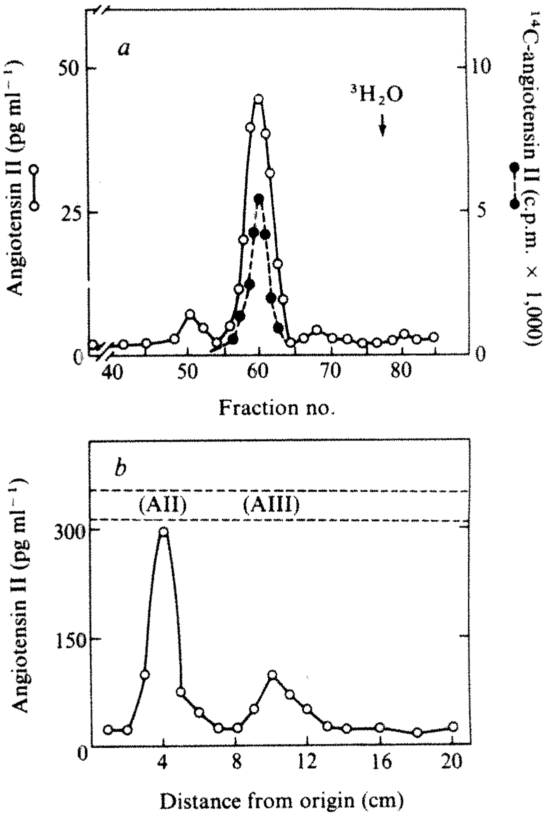


Fig. 4 *a*, Chromatography of adrenal capsule extract on Sephadex G-15 (1 × 75 cm) in 0.01 M acetic acid. Angiotensin II concentration was measured by RIA in the individual fractions, and ^{14}C angiotensin II was used as a marker in separate experiments to calibrate the column. *b*, Fractionation of immunoreactive angiotensin II by TLC in butanol–3% ammonia (5 : 1). The positions of angiotensin II and des-Asp^1-angiotensin II are indicated as AII and AIII, respectively.

The identity of the angiotensin-like material in adrenal capsules was examined in acetone extracts of tissue homogenates prepared in 0.01 M acetic acid and heated at 100 °C for 10 min. Chromatography on Sephadex G-15 revealed a single peak of immunoreactive angiotensin II, corresponding to the established elution profile of ^{14}C-angiotensin II. TLC of such adrenal extracts revealed two peaks of immunoreactive material, the major component of about 80% corresponding to the R_F of angiotensin II, and a smaller peak corresponding to des-Asp^1-angiotensin II (Fig. 4). The angiotensin II peak was eluted with methanol–ammonia (2 : 1), dried down and redissolved in phosphate buffer for evaluation of biological activity by adrenal radioligand–receptor assay and by bioassay in collagenase-dispersed adrenal capsular cells[14]. The eluted immunoreactive angiotensin displaced ^{125}I-angiotensin II from adrenal membranes and stimulated aldosterone production in glomerulosa cells, with equivalent potency to the peptide concentration determined by radioimmunoassay.

The persistence of normal circulating levels of immunoreactive angiotensin II after nephrectomy was observed in assay conditions which detect extremely low blood levels of the peptide after converting-enzyme inhibition[4] or sodium loading. As plasma renin in addition to injected angiotensin II is rapidly cleared from circulation, generation of angiotensin II in the tissues must account for the presence of the peptide in the blood after nephrectomy. Such local production was supported by the demonstration of angiotensin II in several tissue extracts, including the adrenal capsule. The possible contribution of angiotensinase-dependent hydrolysis of tracer angiotensin II to the apparent presence of the tissue peptide has been noted in kidney extracts[15,16]. This effect was avoided in the present study by boiling the homogenates for 10 min before acetone extraction. Also, after chromatographic purification, the angiotensin-like material maintained its immunoreactivity and biological activity. The generation of such significant amounts of biologically active angiotensin II in the tissues may account for the normal blood levels of the peptide detected 24 h after nephrectomy. Although decreases in blood angiotensin II levels were previously observed 18 h after nephrectomy in sodium-restricted rats[12], this earlier fall could be related to the higher plasma angiotensinase activity present during sodium deficiency[17].

Renin-like[18–20] and converting-enzyme[21–23] activity are present in several tissues; we have confirmed that renin activity is present in the adrenal capsule and other tissues, and seems to vary with the PRA during altered sodium intake. Such tissue renin could have at least partly accumulated from renal renin in the circulation, as renin activity in the aorta begins to decrease several hours after nephrectomy[24]. In a related study, PRA fell rapidly after nephrectomy, whereas the blood pressure and the hypotensive response to converting-enzyme inhibition decreased over several hours[25]. Such a process would also be consistent with the gradual fall of blood angiotensin II observed in the present experiments, with no major decrease until 48 h after nephrectomy. However, tissue isorenin is probably responsible for the local generation of a significant proportion of the angiotensin II that is present in the tissues and continues to circulate after bilateral nephrectomy. Normal levels of renin and angiotensin II have been found in the blood of nephrectomized humans and did not rise after sodium depletion[26,27]. Such findings, with those presented here, suggest that extra-renal renin may provide a basal level of local and circulating angiotensin II against which the actions of renal renin are superimposed in normal physiological conditions.

Received 2 July; accepted 20 November 1980.

1. Hauger, R. L., Aguilera, G. & Catt, K. J. *Nature* **271,** 176–178 (1978).
2. Aguilera, G., Menard, R. & Catt, K. J. *Endocrinology* **107,** 55–60 (1980).
3. Aguilera, G. & Catt, K. J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4057–4061 (1978).
4. Aguilera, G., Schirar, A., Baukal, A. & Catt, K. J. *Circulation Res.* **46,** Suppl. I, 118–127 (1980).
5. Mukherjee, C., Caron, M. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1945–1949 (1975).
6. Kahn, R. C., Neville, D. M. & Roth, J. *J. biol. Chem.* **248,** 244–250 (1973).
7. Hsueh, A. J. W., Dufau, M. L. & Catt, K. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 592–595 (1977).
8. Pernollet, M. G., Devynck, M. A., Matthews, P. G. & Meyer, P. *Eur. J. Pharmac.* **43,** 361–372 (1977).
9. Glossmann, H., Baukal, A. & Catt, K. J. *J. biol. Chem.* **249,** 825–834 (1974).
10. Menard, J. & Catt, K. J. *Endocrinology* **90,** 422–430 (1972).
11. Cain, M. D., Coghlan, J. P. & Catt, K. J. *Clinica chim. Acta* **39,** 21–34 (1972).
12. Douglas, J. & Catt, K. J. *J. clin. Invest.* **58,** 834–843 (1976).
13. Faden, V. B. & Rodbard, D. *EXPFIT: Exponential Data Processing,* 73 (National Institutes of Health, Maryland, 1975).
14. Aguilera, G. *et al. Endocrinology* **104,** 1279–1285 (1979).
15. Johnston, C. I., Mendelson, F. A. O., Hutchinson, J. S. & Morris, B. in *Mechanisms of Hypertension* (ed. Sambhi, M. P.) 238–252 (Excerpta Medica, Amsterdam, 1973).
16. Morris, B. & Johnston, C. I. *Endocrinology* **100,** 1409–1416 (1977).
17. Leary, W. F. & Ledingham, J. G. *Nature* **227,** 178–179 (1970).
18. Ganten, D., Hayduk, K., Brecht, H. M., Boucher, R. & Genest, J. *Nature* **226,** 551–552 (1970).
19. Ryan, J. W. & Johnson D. C. *Biochim. biophys. Acta* **191,** 386–371 (1969).
20. Weinberger, M. H. *et al. Circulation Res.* **40,** Suppl. I, 1–3 (1977).
21. DiSalvo, J. & Montefusco, C. B. *Am. J. Physiol.* **221,** 1576–1579 (1971).
22. Yang, H. Y. T. & Neff, N. H. *J. Neurochem.* **19,** 2423–2428 (1972).
23. Cushman, D. W. & Chueng, A. W. *Biochim. biophys. Acta* **250,** 261–263 (1971).
24. Thurston, H., Swales, J. D., Bing, R. F., Hurst, E. C. & Marks, E. S. *Hypertension* **1,** 643–649 (1979).
25. Thurston, H. & Swales, J. D. *Clin. Sci. molec. Med.* **52,** 299–304 (1977).
26. Yu, R., Anderton, J., Skinner, S. L. & Best, J. E. *Am. J. Med.* **52,** 707–711 (1972).
27. Deneheffe, J. *et al. Circulation Res.* **39,** 183–190 (1976).

Gating of a muscle K^+ channel and its dependence on the permeating ion species

P. R. Stanfield, Frances M. Ashcroft & T. D. Plant

Department of Physiology, University of Leicester,
Leicester LE1 7RH, UK

In excitable cells, ions permeate the cell membrane through ionic channels, some of which open and close in response to changes in the potential difference across the membrane. It has been supposed that this opening and closing (or gating) process is largely independent of the permeating ion. However, we show here that the gating of the resting potassium permeability of frog skeletal muscle depends on the species of ion which carries current across the membrane. The potassium permeability investigated allows K^+ to move in across the membrane more easily than out[1,2]. This property is known as inward or anomalous rectification and is shared by cell membranes of skeletal muscle[1–3], egg[4] and certain other cells[5–7]. In both egg cells[8] and skeletal muscle fibres[9], the group IIIB metal ion Tl^+, which can replace K^+ in several other systems in experimental conditions[10–12], also permeates the inward rectifier. Indeed, Tl^+ is more permeant than K^+ (refs 8, 9). However, when Tl^+ carries current inwards across the membrane, the inward rectifier inactivates over a brief period when the membrane is hyperpolarized, whereas when K^+ carries current, the permeability increases with time under hyperpolarization.

Our experiments were carried out on frog sartorius muscle in voltage-clamp conditions[13] at 1–3 °C, and Fig. 1 summarizes the main result. In muscle fibres immersed in a solution containing 80 mM K^+, the current flowing under hyperpolarization increases with time (Fig. 1*a*) along an exponential time course whose time constant falls with increasing hyperpolarization[14–16].

Thus, in K^+, steady-state currents are larger than initial currents (Fig. 1*b*), although initial currents themselves yield an inwardly rectifying current–voltage relation[14–16]. In the solution we used there was a subsequent slow fall in current caused by K^+ becoming depleted from the lumen of the transverse tubular system, in whose walls much of the K^+ permeability resides[14,17,18]. (The inactivation of K^+ currents under extreme hyperpolarization, described by Almers[14,19], has been shown to be a voltage-dependent block by Na^+ (ref. 18) and our solutions contained no Na^+.)

In contrast to the results in K^+, when muscle fibres are immersed in a solution containing 80 mM Tl^+ (but no K^+), the currents which flow under hyperpolarization (Fig. 1*c*) inactivate along an exponential time course. Thus, steady-state currents are smaller than instantaneous currents, as may be seen from the current–voltage relations of Fig. 1*d*, and the steady-state current–voltage relation has a region of negative slope.

Steady-state currents, expressed as a fraction of instantaneous currents, are plotted against membrane potential in Fig. 2*a*. The steady-state inactivation curve may be fitted by a Boltzmann relation of the form:

$$p_\infty = \{1 + \exp[-(V - V_p)/a]\}^{-1} \qquad (1)$$

where p_∞ is the fractional current in the steady state at a membrane potential V; $p = 0.5$ at a membrane potential $V_p = -172$ mV, and $a = 48$ mV.

The time constants for inactivation are plotted against membrane potential in Fig. 2*b*, showing that the rate of inactivation falls as hyperpolarization is increased, is at its slowest at about −160 to −180 mV, and then increases again under further hyperpolarization. This voltage dependence of the rate of inactivation, and the fact that similar time constants are obtained whether they are measured from the onset of inactivation under a single hyperpolarization or from recovery from inactivation using a two or a three pulse method, suggests that

0028–0836/81/060509—03$01.00

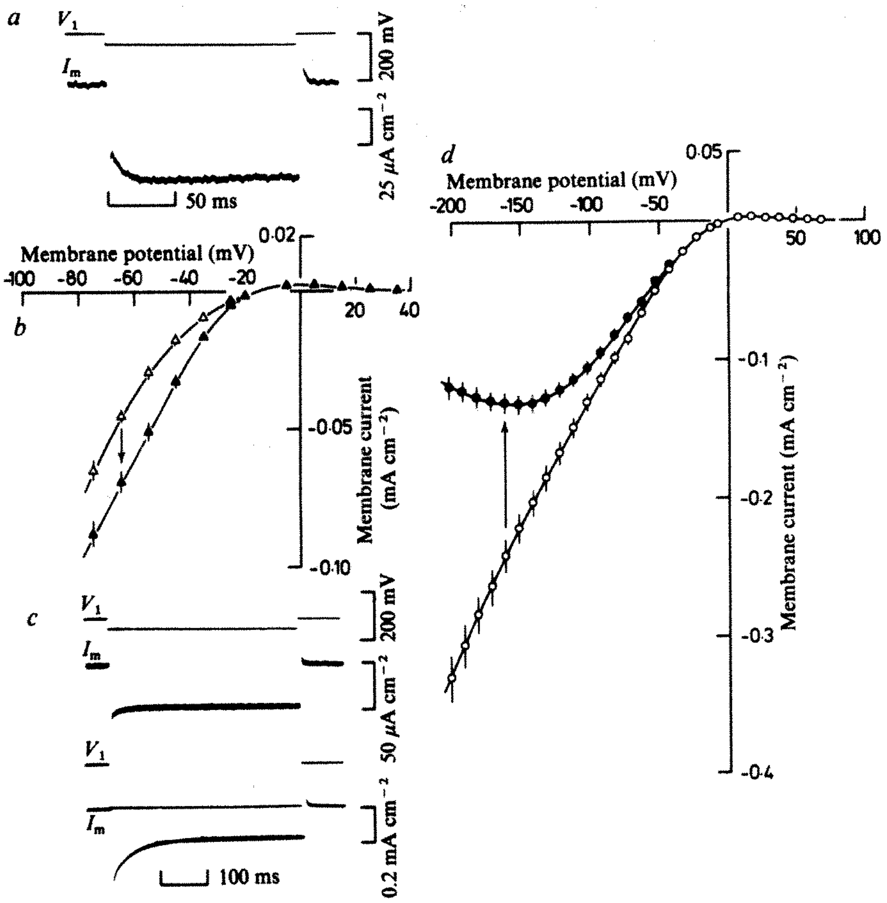


Fig. 1 The gating of inward rectification in K^+ (*a, b*) and in Tl^+ (*c, d*) solutions. *a*, Records of membrane potential (V_1) and membrane current (I_m) from a muscle fibre immersed in a solution containing 40 mM K_2SO_4 (80 mM K^+), 8 mM $CaSO_4$, 113 mM sucrose, 2 mM Tris-maleate buffer (*p*H 7.2). Fibre resting potential (*RP*), −16 mV; holding potential (*HP*), −16 mV, *T*, 1.2 °C, the hyperpolarization is 40 mV in amplitude. *b*, Current–voltage relations for 12 fibres in 80 mM K^+ solution for instantaneous (△) and steady-state (▲) currents. A least squares method was used to fit the exponential turn-on of currents and to compute instantaneous currents by extrapolation. Symbols and vertical bars give mean ±s.e.m. of 12 fibres. *RP*, −14.8 ± 0.6 mV. Abscissa, membrane potential; ordinate, membrane current. *c*, Records of V_1 and I_m from a fibre immersed in a solution containing 40 mM Tl_2SO_4, 8 mM $CaSO_4$, 113 mM sucrose, 2 mM Tris-maleate buffer (*p*H 7.2). Fibre *RP*, −2 mV; *HP*-2 mV; *T*, 1.2 °C, the hyperpolarizations are 40 and 180 mV in amplitude. *d*, Current voltage relations for fibres in 80 mM Tl^+ solutions for instantaneous (○) and steady-state (●) currents. Steady-state currents were computed by correcting for depletion of Tl^+ by extrapolation to zero time using a least squares fit for a straight line to currents between 250 and 400 ms. Depletion is sufficiently slow to allow such a fit at times <400 ms. Instantaneous currents were computed by making a least squares fit to the exponential decline of currents at short times. Abscissa, membrane potential; ordinate, membrane current. *RP*, −2.7 ± 0.4 mV ($n = 9$). In both *b* and *d*, a linear element, treated as leak, has been subtracted from the current voltage relations. This was obtained by extrapolating from currents obtained with depolarizations greater in amplitude than 60 mV through the holding potential.

the inactivation has simple first order kinetics, which could be represented in the following way:

$$\underset{(1-p)}{\text{Channels closed}} \underset{\beta_p}{\overset{\alpha_p}{\rightleftharpoons}} \underset{(p)}{\text{channels open}}$$

with *p* giving the fraction of channels open at a given membrane potential and time, and with the rate constants α_p and β_p approximating to exponential functions of membrane potential (Fig. 2*c*) of a form expected from equation (1).

The rate of inactivation has a high temperature dependence (Fig. 3*a*), with the Q_{10} being 2.75 ± 0.14 ($n = 6$); at room temperature, the onset of the inactivation cannot easily be separated from the capacity transient accompanying hyperpolarization. The voltage dependence of steady-state inactivation is unaltered by changing temperature (Fig. 3*b*). The Q_{10} for the amplitude of the Tl^+ currents was 1.42 ± 0.18 ($n = 6$), similar to that found for K^+ currents[14,19].

The results we describe are very similar to those expected if a blocking cation, present in the external solution, were driven into the membrane by hyperpolarization, blocking the inward movement of Tl^+. However, our solution contained only Ca^{2+},

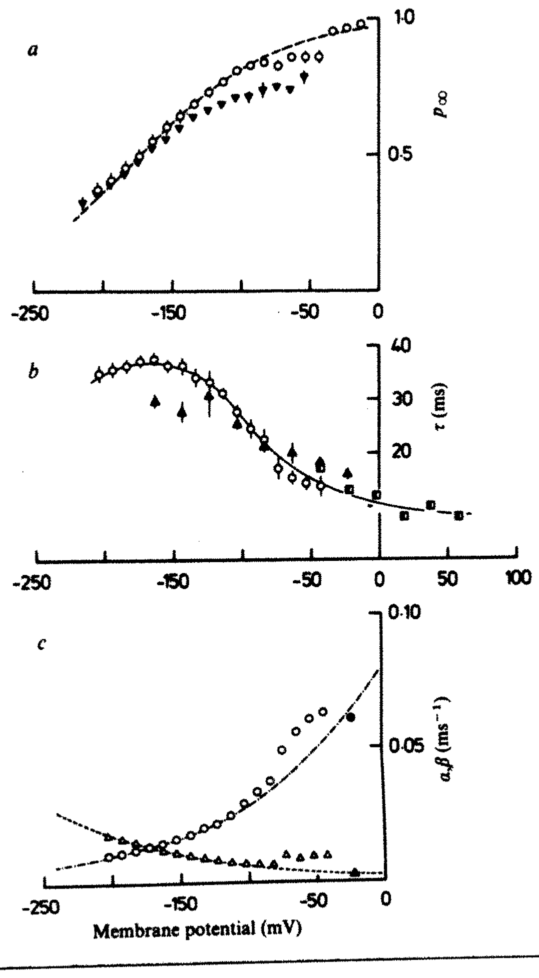


Fig. 2 Inactivation of inward rectification in Tl solutions. *a*, Steady-state inactivation for 80 mM Tl^+ (○) and 40 mM Tl^+ + 40 mM K^+ (▼) solutions. Abscissa, membrane potential; ordinate, p_∞, the steady-state current expressed as a fraction of the instantaneous current. Dashed line, Boltzmann relation given in equation (1) of the text. *b*, Time constants for inactivation (ordinate) plotted against membrane potential (abscissa). ○, Values obtained from inactivation of currents under hyperpolarization (mean ± s.e.m., $n = 9$); ▲, values obtained from recovery of current at the given membrane potentials after a 180 mV hyperpolarization of 150 ms duration (mean ± s.e.m., $n = 4$); ◧, rate of recovery at given membrane potential obtained by applying a hyperpolarizing test pulse of 180 mV amplitude and 150 ms duration a variable interval after an inactivating pulse of 180 mV amplitude and 150 ms duration ($n = 1$). The line to the points was drawn by eye. $T = 1$–3 °C. *c*, Rate constants for the inactivation. ○, ●: α_p; △, ▲: β_p. Open symbols computed from the open symbols of *a* and *b*. Closed symbols, computed from ▲ for a 20-mV hyperpolarization in *b* and the relevant open symbol in *a*. Dashed and dotted line, $\alpha_p = \bar{\alpha}_p \exp[(V - V_p)/2a]$; dashed line, $\beta_p = \bar{\beta}_p \exp[-(V - V_p)/2a]$ with $\bar{\alpha}_p = \bar{\beta}_p = 1.34 \times 10^{-2}\ \text{ms}^{-1}$; $V_p = -172$ mV; $a = 48$ mV.

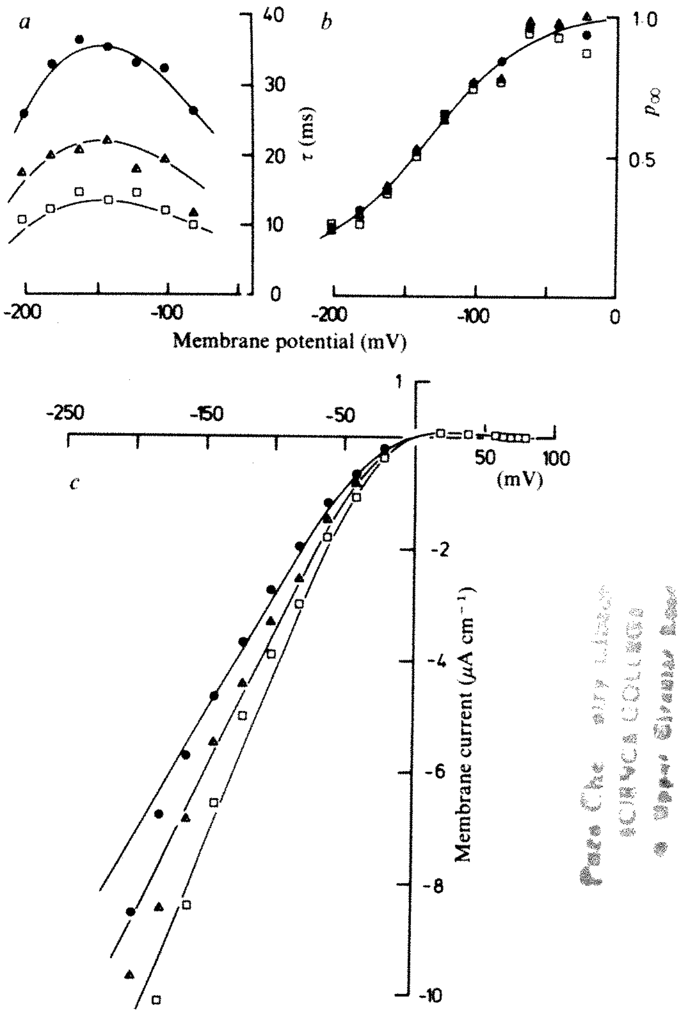


Fig. 3 Q_{10} for inactivation in Tl^+ solutions. *a*, Time constants for inactivation (ordinate) against membrane potential (abscissa), measured at 1.5 °C (●), 5.7 °C (▲) and 10.3 °C (□). The line to the filled symbols (●) was drawn by eye; the others were drawn assuming that the rate of inactivation has a Q_{10} of 3.05 in this fibre. *RP*, −2 mV; *HP*, −2 mV. *b*, Relation between p_∞ (ordinate) and membrane potential (abscissa) at the three temperatures as in *a*. *c*, Relation between membrane current per unit length (ordinate) and membrane potential (abscissa) for the temperature given in *a*. The line to the open symbols (□) was drawn by eye. Those to the other symbols assume a $Q_{10} = 1.57$ for this fibre. Currents per unit length were computed using measured values of r_i (Ω cm^{-1}; sarcoplasmic resistance per unit length) from cable analysis of the response to a 5-mV hyperpolarization (see ref. 13).

apart from Tl^+, and the effect is unlikely to be a block by Ca^{2+} because: its replacement by Mg^{2+} does not alter the inactivation; Ca^{2+} does not block in K^+ solutions[20]; those alkali metal ions (Ba^{2+}, Sr^{2+}) which do block inward rectification do so with an effective valency of 1.4, (ref. 20) not 0.5 ($= aF/RT$) as implied for this inactivation by equation (1).

The inactivation cannot be caused by depletion of Tl^+ from the lumen of the T-system. Depletion would not occur at such a high rate[18], its rate would not have such a high Q_{10} (ref. 19 and Fig. 3*a*), the rate would not fall with increasing hyperpolarization[17,19], and depletion would not give a steady-state current–voltage relation with a region of negative slope[19,21]. The same arguments also exclude the possibility that local intracellular accumulation reduces the driving force on Tl^+. In addition, the steady-state inactivation (Fig. 2*a*) and its rate of onset are little altered in solutions containing 40 mM Tl^+ and 40 mM K^+.

The mechanism of the inactivation we describe is unclear. It is possible that Tl^+ blocks its own inward movement, but such a self-block seems unlikely to account for inactivation time constants as high as 30–40 ms (Fig. 2*b*). An interaction, within the channel, between K^+ from the internal solution and Tl^+ from outside might also be responsible. An alternative is that Tl^+ (but not K^+) stabilizes a conformational change produced by hyperpolarization, which results in closure of the channel. The high Q_{10} for the rate at which inactivation occurs may favour this alternative.

Whatever the mechanism underlying inactivation in Tl solutions, the property we describe—dependence of the gating of an ionic channel on the permeating ionic species—may not be unique to inward rectification. It has recently been shown that the inactivation of the Ca^{2+} channels of certain cells[22–25] depends on the ion (Ca^{2+}, rather than Sr^{2+} or Ba^{2+}) moving through the channels.

We thank Dr N. B. Standen for helpful discussion and critically reading the manuscript, Mr P. Mallory for technical help, the MRC for a fellowship to F.M.A. and the Royal Society for an equipment grant to P.R.S.

Received 15 September; accepted 14 November 1980.

1. Katz, B. *Archs Sci. Physiol.* **3,** 285–299 (1949).
2. Hodgkin, A. L. & Horowicz, P. *J. Physiol., Lond.* **148,** 127–160 (1959).
3. Adrian, R. H. & Freygang, W. H. *J. Physiol., Lond.* **163,** 104–114 (1962).
4. Hagiwara, S. & Takahashi, K. *J. Membrane Biol.* **18,** 61–80 (1974).
5. Hall, A. E., Hutter, O. F. & Noble, D. *J. Physiol., Lond.* **166,** 225–240 (1963).
6. Kandel, E. R. & Tauc, L. *J. Physiol., Lond.* **183,** 287–304 (1966).
7. Nelson, P. G. & Frank, K. *J. Neurophysiol.* **30,** 1097–1113 (1967).
8. Hagiwara, S., Miyazaki, S., Krasne, S. & Ciani, S. *J. gen. Physiol.* **70,** 269–281 (1977).
9. Gay, L. A. thesis, Univ. Leicester (1979).
10. Cavieres, J. D. & Ellory, J. C. *J. Physiol., Lond.* **243,** 243–266 (1974).
11. Hille, B. *J. gen. Physiol.* **61,** 669–689 (1973).
12. Hughes, M. N., Man, W. K. & Whaler, B. C. *J. Physiol., Lond.* **256,** 126P–127P (1976).
13. Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208,** 607–644 (1970).
14. Almers, W. thesis, Univ. Rochester (1971).
15. Hestrin, S. *Soc. Neurosci. Abstr.* **5,** 292 (1979).
16. Hagiwara, S., Miyazaki, S. & Rosenthal, N. P. *J. gen. Physiol.* **65,** 621–638 (1976).
17. Adrian, R. H. & Freygang, W. H. *J. Physiol., Lond.* **163,** 61–103 (1962).
18. Standen, N. B. & Stanfield, P. R. *J. Physiol., Lond.* **294,** 497–520 (1979).
19. Almers, W. *J. Physiol., Lond.* **225,** 33–56, 57–86 (1972).
20. Standen, N. B. & Stanfield, P. R. *J. Physiol., Lond.* **280,** 169–191 (1978).
21. Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208,** 645–668 (1970).
22. Brehm, P. & Eckert, R. *Science* **202,** 1203–1206 (1978).
23. Tillotson, D. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1497–1500 (1979).
24. Ashcroft, F. M. & Stanfield, P. R. *J. Physiol., Lond.* **308,** 36P (1980).
25. Stanfield, P. R. & Ashcroft, F. M. *Proc. IUPS* **14,** 716 (1980).

Cell-free synthesis and processing of multiple precursors to glucagon

Dennis Shields, Thomas G. Warren, Sara E. Roth & Michael J. Brenner

Department of Anatomy, Albert Einstein College of Medicine,
1300 Morris Park Avenue, Bronx, New York 10461

Glucagon, a polypeptide hormone of 29 amino acids, is synthesized in the islets of Langerhans and immunoreactive forms of the molecule have been found in several tissues[1]. Like many other polypeptide hormones, glucagon is synthesized via a larger precursor molecule, proglucagon[2–7]; however, estimates of its size vary considerably and the biosynthetic relationship between some of the putative precursors and authentic secreted glucagon is unclear. Consequently, it was of interest to investigate the primary translation product of glucagon mRNA to relate its size to that of previously described glucagon precursors[5,6]. Here we provide evidence for three distinct immunoreactive preproglucagon molecules, two of which have an apparent molecular weight (MW) of ~14,000 (14K), the third with a MW of ~16,000 (16K). Furthermore, when microsomal membranes were present during translation, the nascent 14K preproglucagon polypeptides were processed to proglucagon with a higher apparent MW of 15,000. In contrast, the nascent 16K preproglucagon was co-translationally processed to a slightly smaller polypeptide. The data indicate that the 14K and 16K preproglucagons undergo different types of post-translational modification.

When poly(A)-containing mRNA isolated from angler fish islets of Langerhans[8] was translated in the wheat-germ cell-free protein synthesizing system, four major translation products were identified[8–10] (Fig. 1, lane 2). Treatment of the products synthesized in the absence of microsomal membranes with a

0028-0836/81/060511—04$01.00

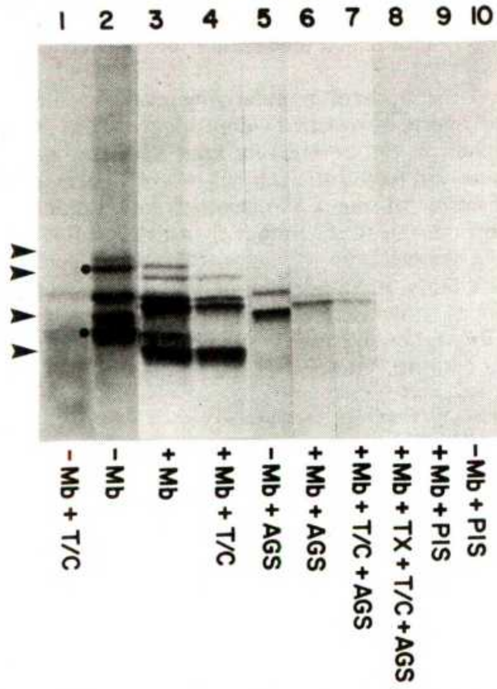


Fig. 1 Synthesis, processing and segregation of proglucagon precursors. Resolution of the translation products was by SDS-polyacrylamide gel electrophoresis using gradients of 12–20% acrylamide[8]; the autoradiograph of the dried gel is shown. Islet mRNA (1 A_{260} unit per ml) was translated in the wheat-germ system, containing 300 µCi ml^{-1} of ^{35}S-methionine, in the absence (−Mb) or presence (+Mb) of dog pancreas microsomal membranes at 4 A_{260} units per ml (ref. 21). Incorporation into acid-precipitable radioactivity was 100,000 and 55,000 µl^{-1} in the absence and presence of membranes, respectively. After 60 min incubation at 26 °C, aliquots of the translation products were prepared for electrophoresis directly (lanes 2, 3) and the remainder was subjected to various post-translational assays. Segregation (assayed by resistance to proteolysis), lanes 1, 4, 7, 8: aliquots were adjusted to 5 mM $CaCl_2$ followed by treatment with a mixture of 250 µg each of trypsin and chymotrypsin per ml (+T/C) for 90 min at 4 °C. Immunoreactive translation products. Lanes 5–8: samples were boiled in 1% SDS then 4 vol of solution A (ref. 22) were added followed by 5 µl of rabbit anti-bovine glucagon serum (+AGS). After incubation for 16 h at 4 °C, the immunoprecipitates were processed for gel electrophoresis[22]. Lane 7: proteolysis, before immunoprecipitation, was terminated by treatment with 1,000 units of Trasylol and 2 mM phenylmethylsulphonyl fluoride; the samples were then treated with SDS followed by solution A[22] and 5 µl anti-glucagon serum. Lane 8: treatment was as for lane 7 except that proteolysis was performed in the presence of 1% Triton X-100 (+TX). Lanes 9, 10: translation products treated with 5 µl of preimmune rabbit serum (+PIS). Upper and lower dots represent the migration of preprosomatostatin and preproinsulin, respectively. Arrows indicate the mobility of soybean trypsin inhibitor, sperm whale myoglobin, egg white lysozyme and bovine proinsulin representing MWs 21,400, 17,000, 14,300 and 8,700 respectively.

Table 1 Relative efficiency of processing and segregation of proglucagons, prosomatostatin and proinsulin

Prehormone	Prepro-hormone	Processed prohormone	Segregated prohormone	% Pro-hormone segregated	% Total pre-hormone segregated
Insulin	8,330	5,250 (63%)	4,413	84	53
Somatostatin	2,630	1,150 (44%)	675	59	26
Glucagon, 14K	2,070	1,990 (96%)	1,360	68	66
Glucagon, 16K	2,360	2,390 (100%)	1,330	56	56

The indicated polypeptides were located by autoradiography, excised from a dried gel similar to that shown in Fig. 1 and the radioactivity (c.p.m.) determined in each gel band[10]. Segregated prohormones are those polypeptides that were resistant to post-translational proteolysis by a mixture of trypsin and chymotrypsin (Fig. 1). Values in parentheses represent the percentage of the total preprohormone processed to the propolypeptide. Note that the radioactivity in the processed 16K proglucagon may be due in part to several unidentified polypeptides in addition to glucagon precursors.

specific rabbit antiserum raised against bovine glucagon, resulted in precipitation of three polypeptides (Fig. 1, lane 5; Fig. 2). Two of the immunoprecipitates, of apparent MW 14,000, were the most prominent and seemed to migrate as a doublet on SDS-polyacrylamide gels. The third immunoreactive polypeptide, which was less abundant than the 14K precursors, had an apparent MW of 16,000. Identical results were obtained using purified monospecific glucagon antibodies (data not shown). As expected, treatment of the translation products with preimmune serum (lane 10) did not result in precipitation of any detectable polypeptides. The three glucagon-immunoreactive translation products have been tentatively designated preproglucagons.

Translation of the islet mRNA in the wheat-germ system supplemented with dog pancreas microsomal membranes, resulted in the synthesis of several additional polypeptides (Fig. 1, lane 3) two of which have been identified as proinsulin[8] and prosomatostatin[9]. Treatment of these cell-free products with anti-glucagon serum resulted in precipitation of two polypeptides (lane 6). The 16K preproglucagon seemed to have been co-translationally processed to a slightly smaller polypeptide which migrated marginally faster than the precursor. In contrast, the microsomal membranes processed both 14K preproglucagons to a higher molecular weight form, ~15K, as judged by the relative intensities of the immunoprecipitates (Fig. 1, lanes 6, 7). Following processing, we were unable to resolve each of these segregated proglucagon molecules as separate bands, and conclude that they have an identical increase in MW. The processed proglucagon molecules were segregated into the heterologous microsomal membrane vesicles, as evinced by their resistance to post-translational proteolysis (lane 7), and consistent with this observation was the finding that proteolysis in the presence of detergent resulted in complete degradation of the segregated proglucagon molecules (lane 8). We conclude that all three nascent proglucagon polypeptides were segregated into the microsomal vesicles but underwent different post-translational processing. The relative efficiency of processing and segregation of the nascent proglucagons as well as proinsulin and prosomatostatin is summarized in Table 1.

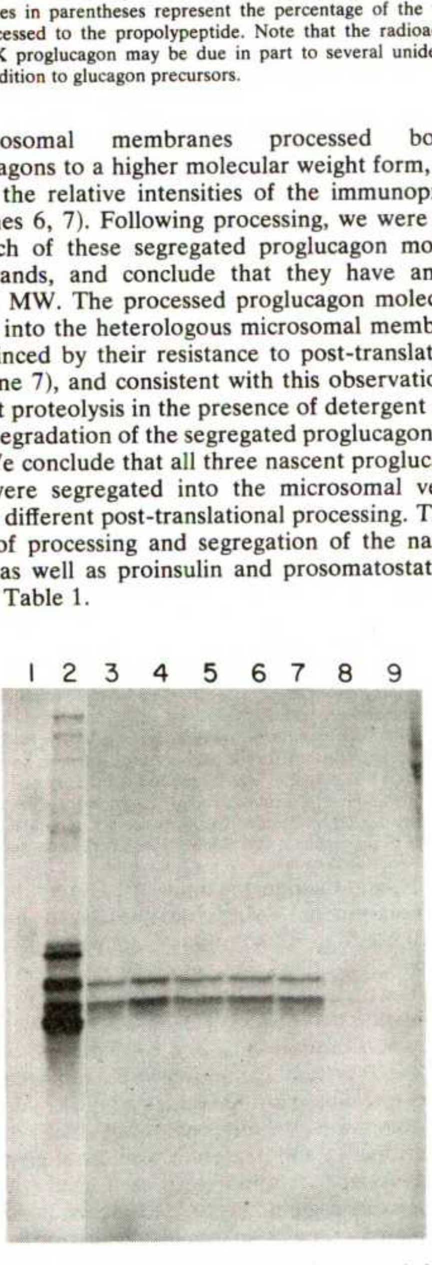


Fig. 2 Immunological characterization of the cell-free translation products having glucagon-like immunoreactivity: competition between unlabelled glucagon and the three putative preproglucagons for binding by anti-glucagon antibodies. Samples of ~300,000 c.p.m. of ^{35}S-methionine-labelled translation products were processed for antibody precipitation as described elsewhere[22], using the following conditions. Lane 1: 5 µl of preimmune serum; lane 2: control total translation products; lane 3: 5 µl of anti-glucagon serum; lanes 4–9: samples were treated with 5 µl anti-glucagon serum in the presence of 0.01 ng, 0.1 ng, 1 ng, 10 ng, 100 ng and 1 µg of unlabelled glucagon, respectively. The immunoprecipitates were processed for gel electrophoresis and autoradiography as outlined in Fig. 1. legend.

In the absence of purified glucagon from angler fish and because the amino acid sequence of angler fish glucagon has not been determined, we could not use heterologous glucagon in peptide mapping experiments to identify the translation products. Instead, we used antibody competition experiments to characterize the glucagon immunoreactive polypeptides further. Increasing concentrations of unlabelled glucagon were added to the cell-free translation products, which were then treated with the rabbit anti-glucagon serum (Fig. 2). With increasing concentrations of unlabelled glucagon there was a diminution in the binding of antibodies to all three preproglucagon polypeptides. Furthermore, all three precursors exhibited parallel responses to competition with unlabelled glucagon (Fig. 3). At the higher concentrations of unlabelled glucagon, antibody binding to all three glucagon precursors was abolished (Fig. 3). Most significantly, the antibody displacement curves of the 14K preproglucagons and the 16K precursor were identical to each other and to that obtained for radioimmunoassay of mature secreted bovine glucagon. These results suggest that the same glucagon antigenic determinants are recognized in all three glucagon precursors. It might be argued that these three molecules were immunoreactive because the antiserum was not specific for glucagon, and contained antibodies against other islet polypeptides. Several lines of evidence make this unlikely. (1) All three cell-free antigens manifested the same displacement curves during antibody competition; it is unlikely that three totally unrelated antigens would exhibit the same competition curves. (2) There was no detectable cross-reactivity of our antiserum with bovine pancreatic polypeptide, neither were the major translation products preproinsulin or preprosomatostatin, immunoprecipitable (Fig. 1). Nevertheless, to show unequivocably that all three precursors contain the glucagon amino acid sequence, we are now purifying angler fish glucagon for use in peptide mapping studies.

Three putative preproglucagon molecules are evident on translation of angler fish islet mRNA in the wheat-germ system, and similar results were obtained using the rabbit reticulocyte cell-free system. Two of the precursor molecules (MW ~14,000) are closely related, at least with respect to having similar migrations on SDS-polyacrylamide gels. In addition, these molecules have the expected molecular weight for proglucagon of 12,000 (ref. 5), possessing an NH_2-terminal signal peptide of 20–30 amino acids[8,9,11]. The third preproglucagon, of MW 16,000, migrated quite separately from the 14K precursors on SDS-gel electrophoresis. The finding of multiple preproglucagon molecules is not without precedent, for two forms of islet preprosomatostatin have recently been described[10,12,13] and two distinct somatostatin molecules have been isolated from catfish islets[14] and sheep hypothalamus[15].

All the glucagon precursor molecules are somewhat smaller than predicted by Patzelt *et al.*[6]; they found a single proglucagon molecule of MW about 18,000 during pulse-chase experiments using isolated rat islets. The discrepancies in molecular weight may be explained by differences in the gel electrophoresis systems. As the preproglucagon molecules are about four or five times larger than the secreted hormone, it is possible that they contain additional hormone amino acid sequences. It would be of interest to determine if the glicentin amino acid sequence is present in any of the putative preproglucagon molecules. Glicentin, which is a 100-amino acid polypeptide isolated from intestinal mucosa[16], has glucagon immunoreactivity and was recently shown[17] to contain the entire sequence of glucagon.

The most intriguing result of these experiments is that the various nascent preproglucagon molecules undergo different post-translational modifications. The 14K preproglucagons exhibit an apparent increase in MW on SDS-gel electrophoresis when the islet mRNA is translated in the presence of microsomal membranes. Similar observations were made by Patzelt *et al.*[6] in their studies on rat islet proglucagon. By analogy to other secretory and membrane glycoproteins[18,19], this result is consistent with the co-translational removal of a signal peptide and the addition of core-sugars to the growing nascent glycopeptide. In preliminary experiments we were unable to demonstrate binding of these processed polypeptides to several different lectin affinity columns. Furthermore, there is no evidence that mature glucagon is a glycoprotein, although this does not exclude the possibility that another region of the precursor molecule contains sugar residues whose conformation prevented binding to the columns. In contrast to the 14K preproglucagons, the 16K precursor was co-translationally processed and segregated to a slightly lower molecular weight proglucagon molecule, results which suggest the removal of an NH_2-terminal signal peptide[8,11]. Thus, it would be of considerable interest to investigate the different modes of post-translational processing that the various processed and segregated proglucagon molecules undergo, particularly with respect to prohormone conversion to authentic glucagon.

During the preparation of this manuscript, a 14.5K precursor to angler fish proglucagon was reported[20].

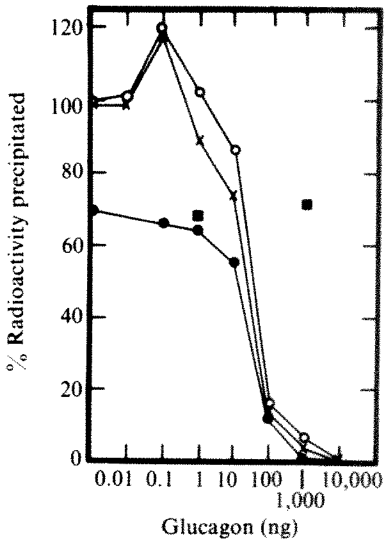


Fig. 3 Quantification of the competition between unlabelled glucagon and preproglucagons for binding by anti-glucagon antibodies. Incubations and treatments were as outlined in Figs 1 and 2 legends. Post-translational incubation with rabbit anti-glucagon serum in the presence of increasing amounts of unlabelled glucagon was as indicated. Control samples received 5 µl of preimmune serum. The bands corresponding to the 16K and both 14K preproglucagons (not separately resolved) were located by autoradiography, excised from the dried gel and solubilized[10] and the radioactivity determined by liquid scintillatin counting. 100% precipitation was equivalent to 1,100 c.p.m. (16K precursor ○—○) and 2,400 c.p.m. (14K preproglucagons ×—×). Background precipitation determined with preimmune serum was 199 c.p.m. Results were calculated after background subtraction. For radioimmunoassay of the anti-glucagon serum (●), aliquots of ^{125}I-glucagon containing 75,000 c.p.m. of trichloroacetic acid-precipitable radioactivity (~0.75 ng glucagon) were treated exactly the same as the translation products[22] and the indicated concentrations of unlabelled glucagon were added followed by 2 µl of anti-glucagon serum. The antibody–antigen complex was precipitated as described previously[9] and the radioactivity determined in a γ counter. ■, Competition between ^{125}I-glucagon and pancreatic polypeptide (10 ng and 1 µg).

We thank Dr R. Chance (Lilly) for gifts of bovine glucagon, pancreatic polypeptide and rabbit antibovine pancreatic polypeptide serum, and Gary Weiss for helpful discussions. This work was supported by NIH research grant AM-21860 to D.S. and a career development award from the Juvenile Diabetes Foundation. T.G.W. is a post-doctoral fellow of training grant AM-07329 and S.E.R. is a predoctoral fellow of the NSF.

Received 19 August; accepted 21 November 1980.

1. Noe, B. D., Fletcher, D. J. & Bauer, G. E. in *Biochemistry, Physiology and Pathology of the Islets of Langerhans* (eds Cooperstein, S. J. & Watkins, D. T.) (Academic, New York, in the press).
2. Rigopoulou, D., Valverde, I., Marco, J., Faloona, G. R. & Unger, R. H. *J. biol. Chem.* **245**, 496–501 (1970).
3. Tager, H. S. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2321–2324 (1973).
4. Noe, B. D. & Bauer, G. E. *Endocrinology* **89**, 625–651 (1971).
5. Noe, B. D. & Bauer, G. E. *Endocrinology* **97**, 868–877 (1975).
6. Patzelt, C., Tager, H. S., Carroll, R. J. & Steiner, D. F. *Nature* **282**, 260–266 (1979).
7. Tager, H. S. & Markese, J. *J. biol. Chem.* **254**, 229–2233 (1979).

8. Shields, D. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2059–2063 (1977).
9. Shields, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4074–4078 (1980).
10. Shields, D. *J. biol. Chem.* **255**, 11625–11629 (1980).
11. Blobel, G. *et al.* in *Secretory Mechanisms Soc. exp. Biol. Symp.* **33** (eds Hopkins, C. R. & Duncan, C. J.) 9–36 (Cambridge University Press, 1979).
12. Goodman, R. H., Lund, P. K., Jacobs, J. W. & Habener, J. F. *J. biol. Chem.* **255**, 6549–6552.
13. Hobart, P., Crawford, R., Shen, L. P., Pictet, R. & Rutter, W. J. *Nature* **288**, 137–141 (1980).
14. Oyama, H., Bradshaw, R. A., Bates, O. J. & Permutt, A. *J. biol. Chem.* **255**, 2251–1154.
15. Bohlen, P. *et al. Biochem. biophys. Res. Commun.* **96**, 725–734 (1980).
16. Jacobsen, H., Demandt, A., Moody, A. J. & Sundby, F. *Biochim. biophys. Acta* **493**, 452–459 (1977).
17. Holst, J. J. *Biochem. J.* **187**, 337–343 (1980).
18. Katz, F. N., Rothman, J. E., Lingappa, V. R., Blobel, G. & Lodish, H. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3278–3282 (1977).
19. Lingappa, V. R., Lingappa, J. R., Prasad, R., Ebner, K. E. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2338–2342 (1978).
20. Lund, P. K., Goodman, R. H., Jacobs, J. W. & Habener, J. F. *Diabetes* **29**, 583–586 (1980).
21. Shields, D. & Blobel, G. *J. biol. Chem.* **253**, 3753–3756 (1978).
22. Lingappa, V. R., Katz, F. N., Lodish, H. F. & Blobel, G. *J. biol. Chem.* **253**, 8667–8670 (1978).

Relationship of glicentin to proglucagon and glucagon in the porcine pancreas

A. J. Moody, J. J. Holst*, L. Thim & S. Lindkaer Jensen*

Novo Research Institute, DK-2880 Bagsvaerd, Denmark

We have previously[1] isolated from porcine small intestine a peptide known as glicentin. The C-terminal portion of glicentin consists of the sequence of glucagon extended at its C terminus by an octapeptide[2,3], and differs slightly from the sequence of a proposed fragment of proglucagon[4]. Glicentin-like material has been demonstrated in the pancreatic A cell[5,6], wherein it is located in the periphery of the secretory granules[7], whereas glucagon is located in the centre of the granules. To study the relationship of glicentin to the biosynthesis of glucagon, we have now investigated the glucagon-like and glicentin-like peptides in extracts[8] and perfusates of the porcine pancreas. Our findings that a peptide with glicentin-like immunoreactivity, and intermediate in size between glicentin and glucagon, is secreted synchronously with glucagon suggest that this glicentin-related peptide is a major cleavage product of proglucagon.

Gel permeation chromatography of 10 extracts of porcine pancreas resolved the glicentin-like immunoreactivity into three peaks (Fig. 1): one (K_d 0.22 ± 0.01; mean ± s.e.m.) with an apparent size similar to that of glicentin; a second (K_d 0.44 ± 0.01), which accounted for the bulk of the immunoreactivity, intermediate in size between glicentin and glucagon; and a third (K_d 0.66 ± 0.01), which eluted in the leading edge of the glucagon peak. The first peak of glicentin-like immunoreactivity (K_d 0.22) was associated with glucagon-like immunoreactivity (GLI), implying that this peak could be similar to intact glicentin. The second (K_d 0.44) was not associated with GLI and could represent a fragment of glicentin which did not contain the sequence of glucagon. The third peak (K_d 0.66) could have possessed glicentin-like and glucagon-like immunoreactivities. The glicentin-related immunoreactivity and GLI recovered from these columns were immunometrically approximately equal.

Isoelectric focusing of an extract of porcine pancreas (Fig. 2) revealed one peak of glicentin-like immunoreactivity with a *p*I close to 4.0. This immunoreactivity was immunometrically equimolar with, but clearly dissociated from, the GLI, which had a *p*I of 6.4, the isoelectric point of porcine glucagon.

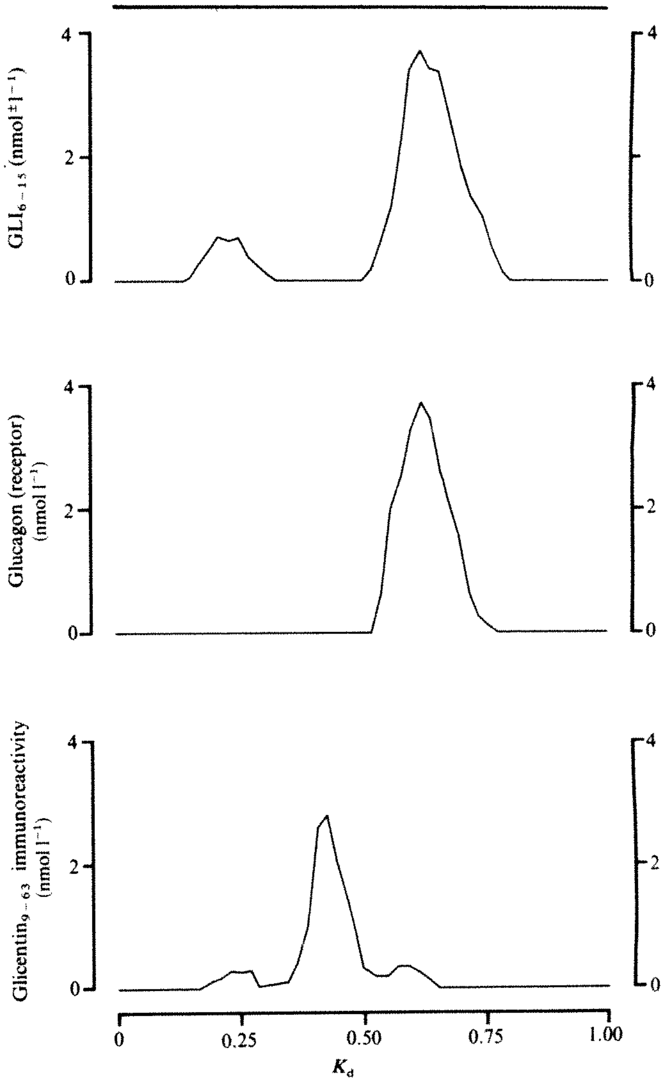


Fig. 1 Gel permeation chromatography of an extract of porcine pancreas on Sephadex G-50. Frozen porcine pancreas was extracted in acid-ethanol at −20 °C, the extract centrifuged and 5 vol diethyl ether added. The peptide-containing aqueous phase was recovered and a sample corresponding to 190 mg of pancreas applied to a 1.6 × 90-cm column of Sephadex G-50 fine equilibrated with 0.5 mol l^{-1} acetic acid. The radio receptor assay for glucagon was carried out using procine liver cell membranes[13]. The glucagon-like immunoreactivities (GLIs) were assayed by the radioimmunoassay for glucagon according to Heding[14]. Anti-glucagon serum 4304 was used to measure GLI homologous with glucagon$_{6-15}$ (ref. 11). Anti-glicentin serum R64 (which reacts with the *N*-terminal portion of glicentin) was used in a radioimmunoassay for glicentin to measure glicentin-like immunoreactivity. The immunoassay consisted of incubating 100-μl aliquots of standards of porcine glicentin, or of unknowns, with equal volumes of antiserum diluted (1:65,000) in a 1:200 dilution of non-immune rabbit serum for 72 h. ^{125}I-glicentin (100 μl of a solution containing 60 pg ^{125}I-glicentin per ml) was then added and the incubation continued for 24 h. The free and antibody-bound ^{125}I-glicentin were separated by centrifugation 24 h after the addition of 100 μl anti-rabbit IgG serum (Dako). In these conditions, R64 does not react significantly with extracts of rat or human pancreas, or with vasoactive intestinal polypeptide, glucagon or gastric inhibitory polypeptide.

To examine further the relationship between porcine pancreatic glicentin-like immunoreactivity and glucagon, we have measured the secretion of these materials by the perfused porcine pancreas. If the glicentin-related peptides and glucagon are located within the same secretory granules—as indicated by the immunocytochemical findings—the secretory patterns of these two materials should be synchronous. Furthermore, if the glicentin-related peptides are cleavage products of glucagon biosynthesis, which are stored and then released together with

---

* Present addresses: Department of Physiology, C, Panum Institute, University of Copenhagen, Copenhagen, Denmark (J.J.H.); Department of Surgical Gastroenterology, Rigshospitalet, Copenhagen, Denmark (S.L.J.).

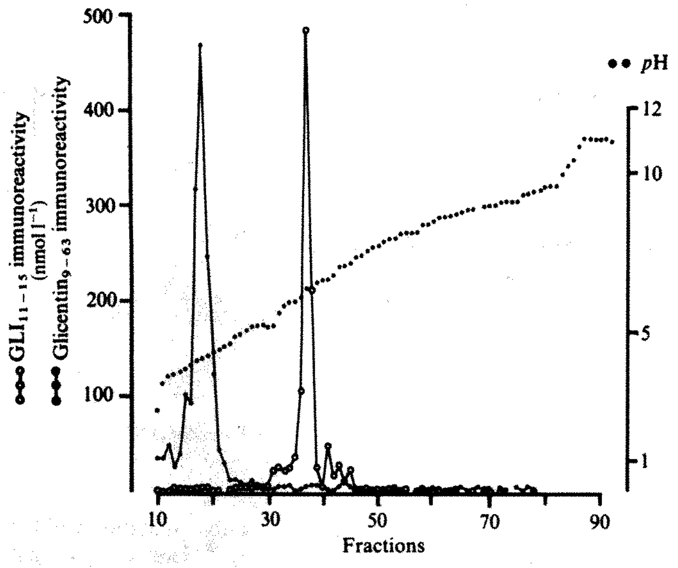


Fig. 2 Isoelectric focusing of an extract of porcine pancreas. A single porcine pancreas was carefully freed from adherent intestine, extracted in acid alcohol and the peptides recovered from the extract as described previously[1]. The peptide pool was desalted on a column of Sephadex G-50 equilibrated with 150 mmol l^{-1} acetic acid. The total peptide pool was recovered from the column and freeze-dried. The freeze-dried material, corresponding to 40 g wet weight of pancreas, was dissolved in ampholine/water and fractionated on a 440-ml LKG isoelectric focusing column with a pH 3.5–10 mixture of ampholines[15] (LKB). The column was emptied in 5-ml fractions, and the glicentin-like immunoreactivity and GLI in the fractions were assayed as described in Fig. 1 legend, except that anti-glucagon serum K 4023 was used to measure GLI homologous with glucagon$_{11-15}$.

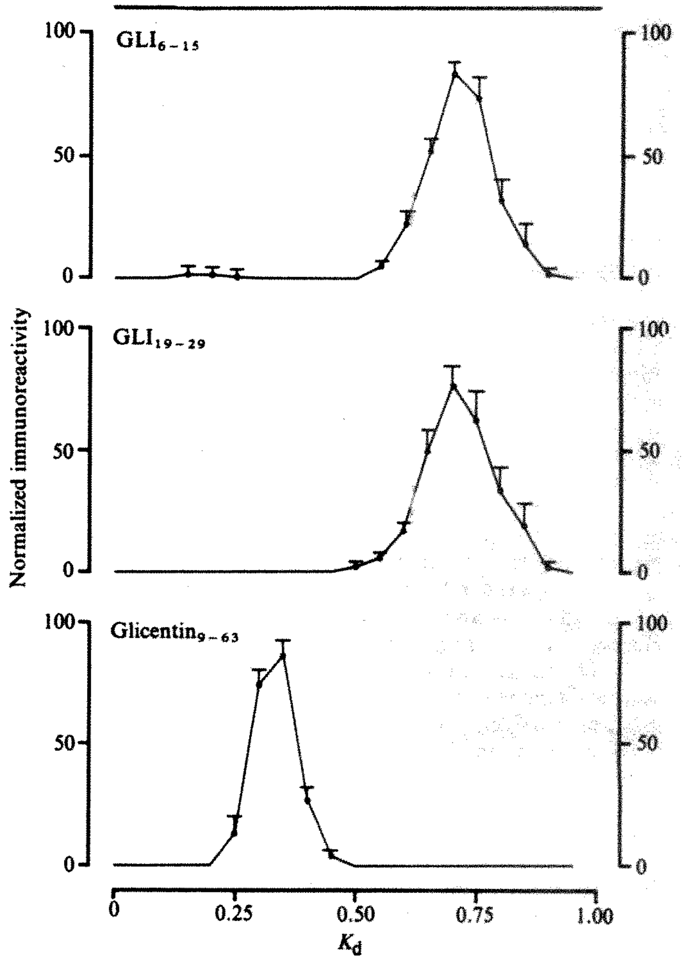
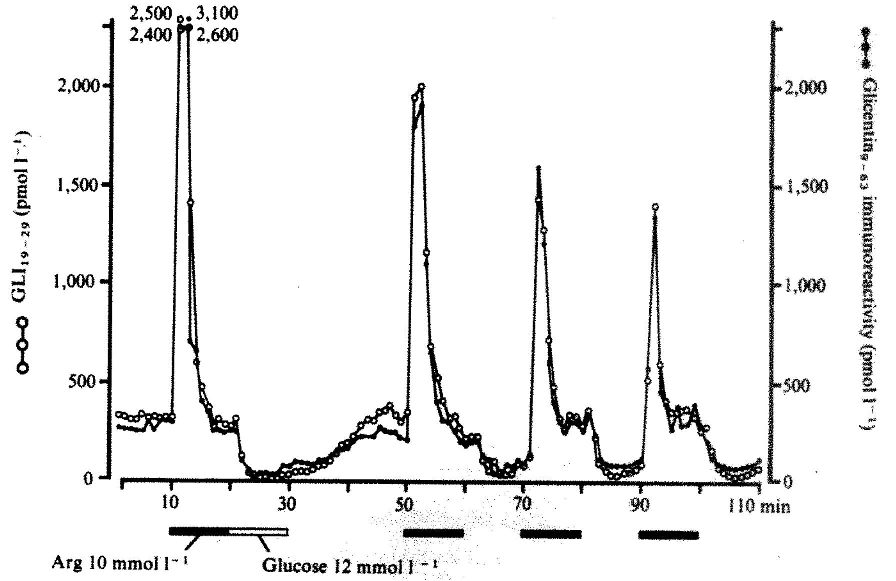


Fig. 4 Gel permeation chromatography of pancreas perfusates on Sephadex G-50 fine. Seven samples of perfusates taken from five different perfused pancreata stimulated with 5 mmol l^{-1} arginine were applied separately to columns of Sephadex G-50 fine equilibrated with 150 mmol l^{-1} NaHCO$_3$ containing 100 mmol NaCl, 20 μg ml^{-1} human serum albumin and 600 μmol l^{-1} thiomersal. The glicentin-like immunoreactivity and GLIs in the effluents were assayed as described previously. The data are the means (+s.e.m.) of the data from seven columns; the relative amounts of immunoreactivities recovered from each column were normalized as follows. The immunoreactivity recovered from each fraction (corresponding to a K_d interval of 0.05) was calculated for each column. The largest amount of immunoreactivity (irrespective of the assay used) recovered in any fraction from the column was taken to be 100, and the immunoreactivities in the other fractions for the column converted to a percentage of this value.

Fig. 3 The effects of arginine and glucose on the secretion of glicentin-related peptides and glucagon by the perfused isolated porcine pancreas. A porcine pancreas was isolated and perfused as described previously[16]. Glucagon release was stimulated by infusion of arginine into the arterial catheter to give a final concentration of 10 mmol l^{-1} and suppressed by the infusion of glucose (12 mmol l^{-1}). GLI and glicentin-like immunoreactivity in the perfusate were assayed as previously described, except that anti-glucagon serum K 4305 was used to measure GLI homologous with glucagon$_{19-29}$ (ref. 11).



glucagon, glicentin-related peptides and glucagon should be secreted in approximately equimolar amounts. Both these suppositions have been found to be correct: stimulation of the glucagon secretion of the perfused porcine pancreas with 10 mmol l^{-1} arginine caused synchronous and immunometrically equivalent increases in the secretion of glicentin-related immunoreactivity and of glucagon (Fig. 3). Gel permeation chromatography of perfusates from pancreases stimulated with arginine showed that the perfusates contained only one glicentin-related peptide, which was larger than glucagon (measured with anti-glucagon sera that reacted with the N- and C-terminal portions of glucagon) and which was devoid of GLI (Fig. 4). Isoelectric focusing of perfusate taken after a stimulation with arginine showed that the glicentin-related material had a *p*I of 4.0. The glicentin-related peptide secreted by the perfused pancreas was therefore similar to the major component of the glicentin-related peptides extracted from the pancreas.

The pancreatic glicentin-related peptides described here can be related to the biosynthesis of glucagon as follows. The peptide with the largest molecular weight, and which was found in small amounts in the pancreas, was very similar to glicentin and could be a stable early intermediate in the biosynthesis of glucagon, that is, proglucagon. The peptide which formed the bulk of the pancreatic glicentin-related material, and which was secreted synchronously with glucagon, could be a major fragment of proglucagon formed during the excision of glucagon from proglucagon. The existence of such a major fragment of proglucagon which is devoid of GLI has been described in rat islets[9].

Final proof of this proposal requires further knowledge of the structure of pancreatic glicentin-related peptides, and a direct demonstration of their relationship to the biosynthesis of glucagon. Supporting evidence for this proposal, which suggests that the biosynthesis of gut GLIs in the intestine and of glucagon in the pancreas share common precursors, is provided by the findings that identical glucagon-containing peptides are found in the rabbit intestine and pancreas[10], that glicentin can be converted into glucagon$_{1-17}$ and glucagon$_{19-29}$ by trypsin and carboxypeptidase B[11], and that ileal glicentin-containing 'L' cells develop glucagon C-terminal immunoreactivity after exposure to the same enzymes[12].

The findings that the pancreas stores and secretes a glicentin-related peptide, which is a possible fragment of proglucagon, suggest that the pancreatic A cell, like the B cell, stores and secretes a large fragment of the precursor molecule of its hormonal secretory product. Further investigation of the nature of pancreatic glicentin-related peptide is required to establish the biological significance of its co-secretion with glucagon.

We thank Karin D. Jørgensen for providing fresh porcine pancreas, and K. Christensen, D. Petersen, W. Jørgensen and E. Gammelgaard for technical assistance.

Received 10 September; accepted 21 November 1980.

1. Sundby, F., Jacobsen, H. & Moody, A. J. *Horm. Metab. Res.* **8,** 266–271 (1976).
2. Moody, A. J., Frandsen, E., Jacobsen, H. & Sundby, F. in *Hormone Receptors in Digestion and Nutrition* (eds Rosselin, G., Fromageot, P. & Bonfils, S.) 55–64 (Elsevier, Amsterdam, 1979).
3. Jacobsen, H., Demandt, A., Moody, A. J. & Sundby, F. *Biochim. biophys. Acta* **493,** 452-459 (1977).
4. Tager, H. S. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **70,** 2321–2324 (1973).
5. Ravazzola, M. *et al. Endocrinology* **105,** 499–508 (1979).
6. Larsson, L.-I. & Moody, A. J. *J. Histochem. Cytochem.* (in the press).
7. Ravazzola, M. & Orci. L. *Nature* **284,** 66–67 (1980).
8. Moody, A. J., Jacobsen, H. & Sundby, F. in *Gut Hormones* (ed. Bloom, S. R.) 367–378 (Churchill Livingston, Edinburgh, 1978).
9. Patzelt, C., Tager, H. S., Carroll, R. J. & Steiner, D. F. *Nature* **282,** 260–266 (1979).
10. Tager, H. S. & Markese, J. *J. biol. Chem.* **254,** 2229–2233 (1979).
11. Holst, J. J. *Biochem. J.* **187,** 337–343 (1980).
12. Ravazzola, M. & Orci, L. *Diabetes* **29,** 156–158 (1980).
13. Holst, J. J. *Diabetologia* **13,** 159–169 (1977).
14. Heding, L. G. *Diabetologia* **7,** 10–19 (1971).
15. Vesterberg, O. & Svensson, H. *Acta chem. scand.* **20,** 820–834 (1966).
16. Lindkaer Jensen, S. *et al. Am. J. Physiol.* **235,** E381–E386 (1978).

0028–0836/81/060516—03$01.00

Molecular cloning and organization of two leghaemoglobin genomic sequences of soybean

D. Sullivan, N. Brisson, B. Goodchild & D. P. S. Verma†

Department of Biology, McGill University, Montreal, Quebec, Canada F3A 2T5

D. Y. Thomas

Department of Molecular Genetics, National Research Council, Ottawa, Canada K1A 0R6

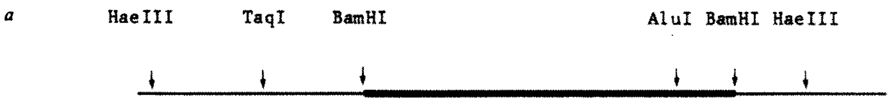
The leghaemoglobins (Lb) are myoglobin-like proteins found in all nitrogen-fixing root nodules of legumes[1–3]. They are encoded by plant nuclear genes[4] which are specifically induced and form the predominant protein in nodules developed in symbiosis with the appropriate species of *Rhizobium*. The Lb is located in the host-cell cytoplasm of the infected cell[5] and is thought to facilitate oxygen diffusion[6,7]. Amino acid sequencing of the soybean Lbs has revealed at least four primary structures differing only in a few amino acids[8–10]. We have previously estimated about 40 copies of Lb sequences in the soybean (*Glycine max* L.) genome by cDNA hybridization[4]. To investigate Lb gene organization and function, we prepared and characterized a Lb cDNA recombinant molecule, pLb1, and used it to isolate two genomic Lb sequences from a library constructed in Charon 4. We report here that the organization of the two genomic Lb sequences is quite distinct and one of them seems to have an intervening sequence(s). Hybridization of pLb1 with genomic DNA from various tissues showed that Lb sequences are dispersed through more than 30 kilobases of genomic DNA and that there is no apparent sequence rearrangement or methylation changes following induction of Lb genes.

Double-stranded cDNA prepared to nodule RNA was digested with *Sau*3A restriction endonuclease and ligated into the *Bam*HI site of the plasmid pBR322 (see Fig. 1 legend). Clones containing Lb sequences were identified by their ability to hybridize with soybean root nodule cDNA, by their hybridization to 9S RNA from root nodules and lack of hybridization to RNA from roots, and by their specific hybridization to sequences in root nodule RNA which are translated into Lb protein in a wheat-germ protein synthesis system (data not shown).

Restriction enzyme analysis of one of the Lb clones demonstrated that the insert was about 145 base pairs long, contained a site for *Alu*I (but not for *Hae*III, *Taq*I or *Hin*fI) and was bounded by sites for *Bam*HI. (The appropriate bases required to regenerate the *Bam*HI site at each end of the insert must have been adjacent to the *Sau*3A sites in the cDNA.) The *Alu*I site corresponds to the position of an *Alu*I site found in a previously described Lb cDNA clone[11]. The DNA sequence of the insert and the predicted amino acid sequence are shown in Fig. 1, together with a partial amino acid sequence of the soybean Lb c$_2$, c$_1$ and Lba. The sequence of the cloned DNA encodes from amino acid residue 99 to the carboxy-terminal phenylalanine (residue 143) of Lbc$_2$ with the exception of an additional valine between residues 102 and 103, and the substitutions of *Asp* for *Asn* and *Gln* for *Glu* (positions 99 and 101 of the Lbc$_2$ sequence[8], respectively). The amino acid sequences of Lbc$_1$ (Fig. 1*e*) and Lba (Fig. 1*f*) differ in 4–8 amino acid residues at this end of the molecule. Thus the cloned Lb sequence represents either a variant of Lbc$_2$, for example, Lbc$_3$ (see ref. 12) or the minor differences may have arisen during molecular cloning[13].

---

† To whom correspondence should be addressed at: Department of Biology, 1205 Avenue Docteur Penfield, Montreal, Quebec, Canada H3A 1B1.



b	GAT	CCT	CAG	TTC	GTG	GTG	GTT	AAA	GAA	GCA	CTG	CTG	AAA	ACA	ATA	AAG	GAG	GCA	GTT	GGG	GAC	AAA	TGG
c	Asp	Pro	Gln	Phe	Val	Val	Val	Lys	Glu	Ala	Leu	Leu	Lys	Thr	Ile	Lys	Glu	Ala	Val	Gly	Asp	Lys	Trp
d	Asn	Pro	Glu	Phe	*	Val	Val	Lys	Glu	Ala	Leu	Leu	Lys	Thr	Ile	Lys	Glu	Ala	Val	Gly	Asp	Lys	Trp
e	Asn	Pro	Glu	Phe	*	Val	Val	Lys	Glu	Ala	Leu	Leu	Lys	Thr	Ile	Lys	Ala	Ser	Val	Gly	Asp	Gln	Trp
f	Asn	Pro	Glu	Phe	*	Val	Val	Lys	Glu	Ala	Leu	Leu	Lys	Thr	Ile	Lys	Ala	Ala	Val	Gly	Asp	Lys	Trp

b	AGT	GAT	GAA	TTG	AGC	AGT	GCT	TGG	GAA	GTA	GCC	TAT	GAT	GAA	TTG	GCA	GCA	GCT	ATT	AAG	AAG	GCA	TTT	TAG	GATC
c	Ser	Asp	Glu	Leu	Ser	Ser	Ala	Trp	Glu	Val	Ala	Tyr	Asp	Glu	Leu	Ala	Ala	Ala	Ile	Lys	Lys	Ala	Phe	COOH	
d	Ser	Asp	Glu	Leu	Ser	Ser	Ala	Trp	Glu	Val	Ala	Tyr	Asp	Glu	Leu	Ala	Ala	Ala	Ile	Lys	Lys	Ala	Phe	COOH	
e	x	Asp	Glu	Leu	Ser	Gly	Ala	Trp	Glu	Val	Ala	Tyr	Asp	Glu	Leu	Ala	Ala	Ala	Ile	Lys	Ala	Lys	COOH		
f	Ser	Asp	Glu	Leu	Ser	Arg	Ala	Trp	Glu	Val	Ala	Tyr	Asp	Glu	Leu	Ala	Ala	Ala	Ile	Lys	Ala	Lys	COOH		

Fig. 1 Characterization of the double-stranded cDNA clone, pLb1. *a*, Diagram of pLb1 insert area; heavy line is cDNA sequence, light lines are pBR322 sequences. *b*, Nucleotide sequence of the 145-base pair insert. *c*, Predicted amino acid sequence encoded by *b*. *d*, *e* and *f*, amino acid sequences of soybean Lb c_2, c_1 and a, respectively[8–10]. The amino acids underlined in *e* and *f* are different from those in *d*. cDNA and double-stranded cDNA[16] were synthesized with reverse transcriptase from oligo (dT)-bound RNA isolated from 3-week-old soybean root nodules[17] (*Glycine max* var. Prize infected with *Rhizobium japonicum* strain 61A76). *Sau*3A-cleaved double-stranded cDNA (10–50 ng) was ligated to *Bam*HI-cleaved pBR322 (1 μg) in 5 μl of 20 mM Tris HCl (*p*H 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.5 mM ATP and 0.1U T_4 DNA ligase (BRL) at 10 °C for 12 h. *Escherichia coli* strain 490 or HB 101 cells were transformed[18] and ampicillin (25 μg ml^{-1})-resistant and tetracycline (20 μg ml^{-1})-sensitive colonies were selected by replica plating on media containing the appropriate antibiotic. Colonies were screened for leghaemoglobin sequences by hybridizing replica filters[19] with ^{32}P-labelled cDNA in 0.6 M NaCl, 50 mM HEPES (*p*H 7.8), 100 μg ml^{-1} heat-denatured sonicated salmon sperm DNA, 0.5% SDS and 50% formamide at 37 °C. Further characterization of positive clones is discussed in the text. The pLb1 insert, including pBR322 sequences, bounded by *Hae*III sites, was isolated by polyacrylamide gel electrophoresis, cut with *Taq*I, and sequenced by the methods of Maxam and Gilbert[20].

Hybridization of restriction enzyme-digested genomic DNA with nick-translated pLb1 reveals several bands (Fig. 2*a*) consistent with the idea of a family of genes[4]. The hybridizing sequences of pLb1 are found within at least 10 regions of the soybean genome differing in their *Eco*RI sites (Fig. 2*a*, tracks 1 and 2). A similar diversity of bands is seen in *Hha*I- (Fig. 2*a*, tracks 3 and 4) and *Hin*dIII-cleaved (Fig. 2*a*, tracks 5 and 6) genomic DNA. The total amount of DNA contained in these fragments ranges from about 30 kilobases (*Hha*I fragments) to almost 50 kilobases (*Eco*RI fragments). The more intensely hybridizing bands (for example, the 1.4- and 4.2-kilobase bands in tracks 1 and 2) may be explained by multiple pLb1 sequences within these fragments or by multiple genes containing the same fragment size. Thus the family of Lb genes is dispersed in the soybean genome and is not composed primarily of identical tandem repeats.

The fragment patterns and intensity of the bands are basically identical whether the DNA was isolated from embryos or nodule tissue. This suggests that the Lb genes are not specifically amplified or rearranged in nodule tissue developed as a result of infection by *Rhizobium*[4]. Furthermore, the *Hha*I results (in addition to those of *Msp*I and *Hpa*II, data not shown) indicate that the induction of the Lb genes is not accompanied by major changes in methylation at these sites.

A soybean genomic library constructed in Charon 4 was screened with cDNA and pLb1 and yielded two clones which contain hybridizing *Eco*RI fragments of ~11.5 (Ch4 Gm Lb11) and 4.2 (Ch4 Gm Lb4) kilobases (Fig. 2*b*, lanes 2 and 3). These fragments correspond to the 11.5 and 4.2-kilobase bands observed in *Eco*RI-cleaved genomic DNA. Both clones contain approximately 14 kilobases of genomic DNA but those sequences which hybridize to the Lb probes are within quite distinct restriction-endonuclease domains (Fig. 3). The use of Ch4 Gm Lb11 and Ch4 Gm Lb4 as hybridization probes to *Eco*RI-cleaved genomic DNA reveals bands which correspond to the *Eco*RI bands in the respective Charon clones, as well as two to four additional bands (data not shown). The relative intensities of the bands indicate approximately equimolar amounts of the fragments in the genome and, when compared with the intensity of hybridization of a ribosomal DNA recombinant probe[14], each band seems to be present in fewer than 10 copies per haploid genome.

The general organizational maps of the two clones were constructed by various single and multiple restriction-enzyme digestions followed by blotting and hybridization (Fig. 3*a*, *b*). A more detailed map of Ch4 Gm Lb11 was obtained by sequential hybridization with pLb1 followed by cDNA to reveal both 3′ and 5′ ends of the sequences. pLb1 sequences represent the 3′ end of the coding region of Lb mRNA (Fig. 1); a complete cDNA probe would represent sequences predominantly to the 5′ end of pLb1. Representative data are given in Fig. 2*c*. Tracks 1 and 6 (*Msp*I and *Hae*III) clearly show the presence of two sequences which are hybridized by both pLb1 and cDNA (1.3 and 1.7 kilobases, track 1) and a 1.6-kilobase band hybridized exclusively to cDNA (track 6). Tracks 2–5 and 7–10 show similar data on pLb1 and/or cDNA hybridizations for the enzyme combination *Mbo*II plus *Hae*III. The fragment hybridized exclusively by cDNA is reduced from about 400 base pairs to 300 base pairs. Applying this sequential hybridization strategy with several restriction enzymes has allowed us to construct the detailed map of Ch4 Gm Lb11 shown in Fig. 3*c*. This map is confirmed by the Smith and Birnstiel[26] method of 5′-end labelling of DNA. Note that the separation of the sequences which hybridize exclusively to cDNA from the nearest pLb1 hybridizing sequences is ~1.3 kilobases. Another region containing pLb1 hybridizing sequences is at least 1 kilobase further away. The region hybridized by the cDNA spans up to 5 kilobases. If the sequences in Ch4 Gm Lb11 represent a complete or even a partial gene, there must be at least one intervening sequence within the Lb coding region. Sufficient DNA to encode the entire mRNA is contained within the exclusively cDNA hybridizing and the nearest pLb1 hybridizing regions. The present data do not allow us to limit the positions of cDNA and pLb1 hybridization sequences, particularly those to the right of the

*Msp*I site. With the above organization of Lb gene, the direction of transcription would be from left to right. In comparison with Ch4 Gm Lb11, the Lb4 has a very restricted region of hybridization with both cDNA and pLb1 indicating that it may not have an intervening sequence or it contains a partial gene sequence. The Lb proteins are very similar[8-10] (see Fig. 1*d*, *e*, *f*) and the cDNA and mRNA sequences seem to cross-hybridize fully[4]. Thus, the genomic sequences hybridized to pLb1 are expected to include all the Lb genes, and Lb clones cannot be identified as encoding any specific Lb protein without sequencing the DNA. Finer mapping and sequencing of fragments could also demonstrate the interrelationship between these genes as well as among other isolated Lb genes in the family.

The organization and regulation of the Lb genes is of interest at several levels[15]: (1) the basic molecular biology of a higher plant nuclear gene; (2) the elegantly orchestrated interactions between the plant and the prokaryote which induces these genes, and (3) the significant agricultural importance of the legume/*Rhizobium* association leading to the fixation of atmospheric nitrogen.

We thank Dr D. Housman for helping to construct and manipulate the Charon libraries and Dr R. Goldberg for providing a soybean library. This study was supported by grants from the Rockefeller Foundation and the Natural Sciences and Engineering Research Council of Canada. The recombinant DNA work was carried out in a B-level facility in accordance with the guidelines of the MRC Canada.

Received 8 August; accepted 19 November 1980.

1. Virtanen, A. I., Jorma, J., Lincola, H. & Linnaslmi, A. *Acta chem. scand.* **1,** 90–111 (1947).
2. Graham, D. H. & Parker, C. A. *Aust. J. Sci.* **23,** 231–232 (1961).
3. Falk, J. E., Appleby, C. A. & Porra, R. J. in *SEB Symposium* 13 (ed. Porter, K. H.) 73–86 (Cambridge University Press, 1959).
4. Baulcombe, D. & Verma, D. P. S. *Nucleic Acids Res.* **5,** 4141–4153 (1978).
5. Verma, D. P. S. & Bal, A. K. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3843–3847 (1976).
6. Wittenberg, J. B., Bergersen, F. J., Appleby, C. A. & Turner, G. L. *J. biol. Chem.* **249,** 4057–4066 (1974).
7. Bergersen, F. J. & Turner, G. L. *J. gen. Microbiol.* **89,** 31–47 (1975).
8. Sievers, G., Huhtala, M.-L. & Ellfork, N. *Acta chem. scand.* **B32,** 380–386 (1978).
9. Nicola, N. A. thesis, Univ. Melbourne (1975).
10. Ellfolk, N. & Sievers, G. *Acta chem. scand.* **25,** 3335–3534 (1971).
11. Truelsen, E., Gausing, K., Jochimsen, B., Jorgensen, P. & Marcker, K. *Nucleic Acids Res.* **6,** 3061–3072 (1979).
12. Fuchsman, W. & Appleby, C. A. *Biochem. biophys. Acta* **579,** 314–324 (1979).
13. Porter, A. G. *et al. Nature* **282,** 471–477 (1979).
14. Varsanyi-Breiner, A. *et al. Gene* **7,** 317–334 (1979).
15. Verma, D. P. S. in *Molecular Biology of Plant Development* (eds Smith, H. & Grierson, D.) (Blackwell Scientific, Oxford, in the press).
16. Ullrich, A. *et al. Science* **196,** 1313–1319 (1977).
17. Verma, D. P. S., Nash, D. & Schulman, H. *Nature* **251,** 74–77 (1974).
18. Mandel, M. & Higa, A. *J. molec. Biol.* **53,** 159–162 (1970).
19. Grunstein, M. & Hogness, D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
20. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
21. Southern, E. *J. molec. Biol.* **98,** 503–517 (1975).
22. Rigby, P. W. J., Dieckman, M., Rhodes, C. & Berg, R. *J. molec. Biol.* **113,** 237–251 (1977).
23. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
24. Benton, W. D. & Davis, R. W. *Science* **196,** 180–181 (1977).
25. Blattner, F. R. *et al. Science* **202,** 1279–1284 (1979).
26. Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3,** 2387–2390 (1976).

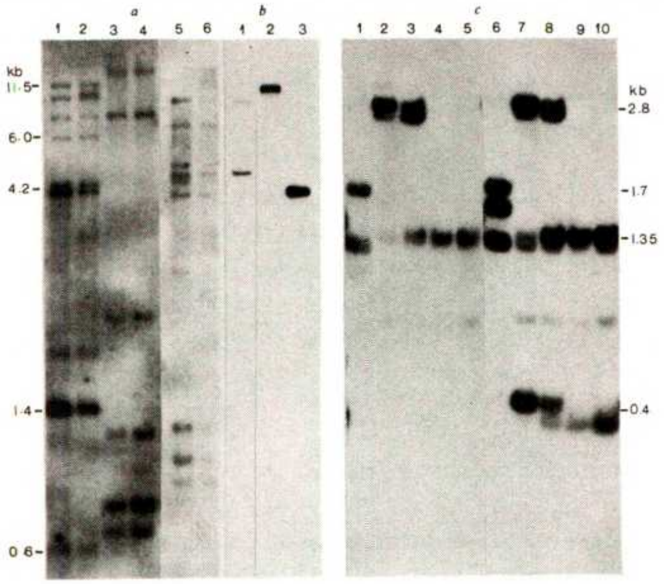


Fig. 2 Analysis of total genomic and cloned genomic Lb sequences. *a*, Restriction endonuclease-digested soybean DNA (5 μg per track) or *b*, *c*, recombinant clone DNA (~0.5 μg per track) was separated by agarose gel electrophoresis, blotted on to nitrocellulose paper[21] and hybridized with ^{32}P nick-translated[22] pLb1 at ~10^8 c.p.m. μg^{-1} and root nodule cDNA (*c*, tracks 6–10) at ~10^7 c.p.m. μg^{-1}. Hybridization conditions were 6×SCP (1×SCP = 0.1 M NaCl, 0.03 M Na_2HPO_4, 0.001 M Na_2 EDTA, *p*H 6.2), 1% Sarkosyl, 2.5% Dextran sulphate at 68 °C for 24–48 h. Exposure of the hybridized filter to pre-fogged X-ray film at −70 °C with intensifying screens was for *a*, 6 d, or *b*, *c*, 12–24 h. Fragment sizes (in kilobases (kb)) were determined from the positions of *Eco*RI-cleaved λ phage and *Taq*I-cleaved pBR322. *a*, Soybean DNA extracted from embryos (track 1) or nodule tissue (track 2) was digested with *Eco*RI; tracks 3 and 4 were digested with *Hha*I; tracks 5 and 6 with *Hin*dIII. *b*, Ch4 Gm Lb11 DNA digested with *Hin*dIII (track 1) or *Eco*RI (track 2); Ch4 Gm Lb4 DNA digested with *Eco*RI (track 3). *c*, Ch4 Gm Lb11 DNA digested with *Hae*III and *Msp*I (tracks 1 and 6) or *Mbo*II followed by *Hae*III partial digestions (tracks 2–4 and 7–9) and complete *Hae*III digestion (tracks 5 and 10). Tracks 1–5 were hybridized to pLb1 and 6–10 to pLb1 followed by Lb cDNA. The genomic recombinant molecules containing Lb sequences were isolated from a library constructed by R. Fisher and R. Goldberg in Charon 4 with partial *Eco*RI-digested soybean DNA by the procedures of Maniatis *et al.*[23]. About 400,000 plaques were screened[24] with ^{32}P-labelled cDNA and pLb1 and three 'positives' were plaque purified[25]. Two of these had identical *Eco*RI and *Hin*dIII sites and have been designated Ch4 Gm Lb11 in reference to the size of the *Eco*RI fragment which is hybridized to pLb1. The third recombinant (Ch4 Gm Lb4) contains a 4.2-kilobase *Eco*RI fragment that hybridized to pLb1.

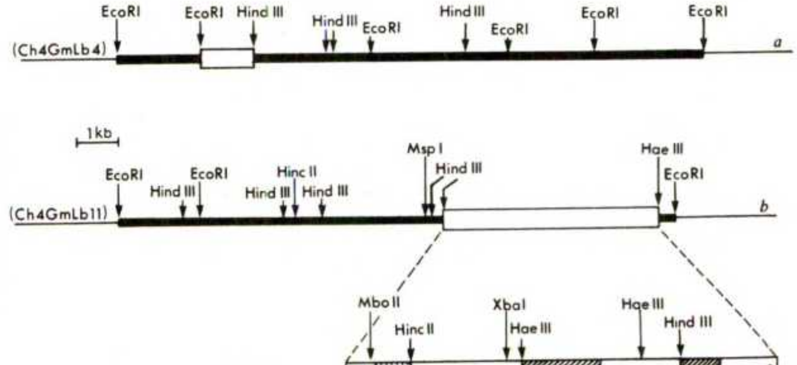


Fig. 3 Organizational map of two Lb genomic sequences. *a*, *b* represent entire insert (heavy line) containing regions of pLb1 hybridization (open box) representing 1.3 and 5 kilobases (kb) in Ch4 Gm Lb4 and Ch4 Gm Lb11, respectively. Light lines represent the arms of Charon 4. *c*, Detailed map of Ch4 Gm Lb11 containing Lb sequences. The dotted area represents sequences exclusively hybridized to Lb-cDNA whereas hatched areas contain sequences hybridized to pLb1.

BOOK REVIEWS

Fast breeders: all cons, no pros?

Alvin M. Weinberg

The Fast Breeder Reactor: Need? Cost? Risk? Edited by Colin Sweet. Pp.232. (Macmillan Press: London, 1980.) £20.

THAT nuclear energy would grind to a halt unless a breeder were developed was understood by the original workers in nuclear energy. What was not clear then, and what is still not entirely clear, is the kind of breeder that ought to be developed. In 1952, Bennett Lewis, father of the Canadian CANDU system, along with E. O. Lawrence, proposed the electric breeder based on an enormous proton accelerator. A variant of electric breeding, the so-called fission/fusion hybrid is now being discussed fairly seriously. I, along with H. G. MacPherson, espoused breeders based on the thorium cycle, particularly those using molten salts. And a thorium breeder based on light-water reactor technology has been built by H. G. Rickover in Shippingport, Pennsylvania. But the plutonium fast breeder, pioneered by W. H. Zinn, gained momentum and is now essentially the only show in town.

We could always give a plausible rationale for the ultimate deployment of breeders, but we could never say when breeders, of whatever kind, would be necessary. Since the fuel cycle cost of the breeder is far less sensitive to the cost of the raw uranium than is the fuel cycle cost of non-breeder pressurized water or gas-cooled reactors, eventually breeders would provide cheaper energy than would non-breeders. Our argument for aggressive development and quick deployment of the breeder was therefore based on the hope that breeders might be cheaper than non-breeders. The only way to find whether or not this was the case was to build large breeders. This remains the strongest argument for vigorous development of a full-size fast breeder. We had two other, albeit secondary, arguments favouring quick deployment of breeders: they would confer on their users a measure of energy autarky and thus would relieve pressure on non-renewable energy sources; and their deployment would put a virtual end to the mining of uranium, the only part of the fuel cycle that poses a technically difficult problem of waste disposal.

The basic economic arguments have hardly changed in the intervening 30 years, as these proceedings of the Polytechnic of the South Bank 1978 Conference on the Fast Breeder demonstrate. The main change is the emergence of non-economic issues — political, social and ethical — that were barely identifiable in the early days. As the editor, Colin Sweet, says,

> It is necessary to subject the official view to a critical assessment and it was the purpose of the conference . . . to do that . . . the issues . . . go far beyond what is to be found in the official literature . . . There is no question of trying . . . to strike 'a balanced view,' because there is no way in which such a phrase can be interpreted.

The resulting 14 papers and five appendices examine not only the old questions (will the fast breeder be economic; and will it be safe?) but also the far broader question of whether nuclear energy itself is needed. Most of the authors are opposed to nuclear energy. Their attack on the fast breeder is an attack on nuclear energy itself. When David Widdicombe argues that nuclear energy and civil liberties are incompatible or when Professor Patricia Lindop says that "the radiation hazard to the population may prove to be the factor limiting the utilization of nuclear energy" they are calling for an end to nuclear energy, not merely an end to the fast breeder.

For those seeking a balanced account of the debate over the fast breeder reactor, these proceedings will be disappointing: it is not only that the positions are far too polarized, it is that there is no real confrontation between the different points of view. Though the nuclear advocates are called to task, their opponents always have the last word. I find those last words, such as Peter Odell's claim of 70% efficiency for oil burning heaters, incomplete and tendentious; they scarcely conceal a profound distrust of nuclear energy.

The editor of *Nature,* in asking me to review this book, implied that the fast breeder reactor issue in the United Kingdom is controversial and thus that the book would be best dealt with by an outsider. As an American, I am reluctant to become involved in a family squabble. Nevertheless, for what it is worth, I offer these points for consideration.

(1) The denial of nuclear energy amounts to a denial of human ingenuity. For example, much is made of the loss of civil liberties entailed in nuclear energy. Yet, since Strategic Air Command bases can be made secure with minimal infringement on civil liberties, I don't see why the same cannot be done for nuclear sites.

(2) I have never been a strong supporter of the fast breeder reactor; I have always thought the thermal system based on thorium was a better bet. In the best of all possible worlds, many different breeders would be pursued, and the choice of the fast breeder would not have pre-empted development of the alternatives to the point where a fair comparison could be made between the various possibilities.

But in the real world, the fast breeder is now practically the only option. Despite the real possibility that fast breeders will prove too costly to make economic sense in the short run, active development of the breeder is at least as sensible as is active development of solar energy. The two are the only inexhaustible energy sources that we know to be technically feasible. We shall know the price of the breeder only by building several, just as we shall know the price of solar power towers only after we have built some.

(3) In the final analysis, we must decide whether the risk in going ahead with the fast breeder and then discovering we do not need it (fusion may work, or solar may be cheap) is less than the risk of not going ahead with the breeder and discovering in, say 2025, that we need it. Considering what is at stake — our energy future, not to speak of the spectre of CO_2induced climate change — I believe we must go ahead with development of the breeder, expensive though this may be.

France, which has chosen to generate 40 per cent of its primary energy from non-fossil sources by 1990, has decided to depend on nuclear energy — that is, on the breeder. It is not unlikely that, in the year 2000, the rest of us may envy the degree of energy autarky enjoyed by France as a result of this decision while we bicker over whether nuclear energy can be mended, and fight over the oil that remains in the Middle East. □

Alvin M. Weinberg is Director of The Institute for Energy Analysis, Oak Ridge Associated Universities, Tennessee.

Conservation: the hopes and the fears

Brian Bertram

Conservation Biology: An Evolutionary–Ecological Perspective. Edited by Michael E. Soulé and Bruce A. Wilcox. Pp.395. (Addison-Wesley/Sinauer: 1980.) Pbk £7.95, $14.50.

THIS will be a disturbing book to anyone who believes that conservation is a matter merely of defining and protecting a number of reserve areas which are representative of ecosystems being destroyed. Michael Soulé and Bruce Wilcox have gathered a fine series of papers demonstrating the vast problems ahead of us all, and showing what we can expect the results of our conservation efforts to be.

Part I contains four chapters on the ecological principles of conservation. They concentrate on tropical forests, which are particularly vulnerable and of which it is likely that two-thirds will have disappeared within a mere 20 years. These early contributions by Gilbert, Eisenberg, Diamond and Foster, show what it is that we risk losing — a fabulously intricate set of interactions among plants and animals, varying both with locality and over time. We are clearly shown that small isolated "representative" portions of the vanishing tropical forests do not contain or cannot maintain more than a fraction of the species of the whole area represented.

Isolated reserves dotted about in a barren waste of altered or destroyed habitat act like islands; the five chapters in Part II consider the consequences of this "insularization" (horrid word). It is well known that true islands generally contain fewer species than the adjacent mainland, and a frequent cause is that the small populations on small islands are highly vulnerable to extinction. Recolonization depends on accessibility from other land areas. Wilcox, and Terborgh and Winter, draw on principles from insular ecology and extrapolate to isolated reserves.

Even if they do not become extinct, small populations face unusual evolutionary problems, as is shown in Franklin's and Soulé's chapters. They are subject to genetic drift, and they contain only a limited amount of genetic variability with which to respond to selection pressures. The level of inbreeding among members of the small population rises, fixing deleterious recessive genes and resulting in higher mortality, lower fecundity and distorting the sex ratio towards surplus males. Two particularly important conclusions emerge. First, that the effective population size must be not less than about 50 individuals to keep inbreeding to a level low enough for the population to survive. Second, that in Soulé's words "significant evolutionary change in most higher organisms is coming to a screeching halt in the tropics".

Part III, on captive propagation and conservation, deals with the similarly small populations which can be kept and bred in zoos. Some species can certainly be preserved in captivity even when extinct in the wild. But, as Conway shows, all the zoos in the USA together contain only enough space for viable populations of about 200 species of mammals. Co-ordination among zoos is required and is growing, but recognition by society of their important role is far too slow. Benirschke and colleagues describe some of the variety of modern aids in captive propagation, and Kleiman surveys the application of sociobiological study to animals in captivity; both enable greater breeding potential to be realized. Successful reintroduction to the wild is the ultimate and difficult objective, and Campbell reviews the numerous problems involved and the successes achieved.

The final section in the book deals with exploitation and strategies of conservation. Coe examines ecological interactions between human beings and wildlife in Africa, and assesses the present and future importance of game hunting, pastoralism, the tsetse fly, tourism, and the prospects for the culling, ranching and domestication of wild species of mammal. With Kenya's population now increasing at 4% per year, major problems are looming fast. Whitmore describes the extent and consequences of the destructive exploitation of tropical forests — some 50 hectares have been destroyed and disappeared while you have been reading this review, and another 30 will have gone before you finish. The direct effect on local species of animal is obvious; the indirect effects on all animals everywhere, because of the climatic changes caused, are unknown.

What we should do is the subject of the last two chapters. Pyle discusses the general management of nature reserves in the developed world. But that is not where most of the problem lies, although many of the causes of the problem do, according to Ehrlich. Human population growth threatens all tropical habitats; but so too does economic growth or the expectation of it. The Amazon forests are being erased only partly to provide living land for burgeoning Brazilians but also to provide timber and beef for North America and Europe. Ehrlich urges us to initiate and follow a more aggressive and wider conservation strategy. We must not be misled by the relatively slow rate of habitat loss in our own countries.

The papers in this book are diverse in character, but they all point very clearly in one direction. The editors summarize it eloquently:

> The green mantle of Earth is now being ravaged and pillaged in a frenzy of exploitation by a mushrooming mass of humans and bulldozers. Never in the 500 million years of terrestrial evolution has this mantle we call the biosphere been under such a savage attack. . . There is no escaping the conclusion that in our lifetimes, this planet will see a suspension, if not an end, to many ecological and evolutionary processes which have been uninterrupted since the beginnings of palaeontological time. We hope it is only a suspension — that the horrible onslaught can be stopped before the regenerative powers of ecosystems are killed. This book is a statement of that optimism. . . This is the challenge of the millenium. For centuries to come, our descendants will damn us or eulogize us, depending on our integrity and the integrity of the green mantle they inherit.

This book is not the uninformed emotional outburst of an eco-freak. It is a collection of the informed statements and pleas of 21 experts. But it is well written, so anyone can read it. And everyone should. □

Brian Bertram is Curator of Mammals at the Zoological Society of London.

Weather and water

William R. Cotton
Kevin R. Knupp

Clouds and Storms: The Behaviour and Effect of Water in the Atmosphere. By F.H. Ludlam. Pp.405 + plates. (Pennsylvania State University Press: 1980.) £24.95, $57.50.

THE late Dr Frank H. Ludlam was a keen observer of nature. His ability to combine and synthesize fragments of data and thereby construct data-consistent conceptual models of weather-phenomena is unmatched in this century. In this book, Dr Ludlam focused on the various forms of water in the atmosphere as the underlying factor in controlling the unique characteristics of many weather systems.

The book is largely devoted to the observed characteristics of the physics and dynamics of convective clouds (Chapters 5,7 and 8). Discussion of topics closely related to cloud processes — radiation, atmospheric optics, atmospheric thermodynamics, atmospheric aerosol physics and cloud microphysics — in the preceding chapters serves as a useful background. A transition chapter (6) compares the timescale and energy of clouds to related atmospheric processes. The ninth and final chapter deals with clouds associated with larger-scale systems.

The book is unique in several respects. First, it contains a copious amount of information, particularly on observed details of convective clouds in Chapters 7 and 8. These two chapters alone contain 397 references to various world-wide studies, and provide an excellent reference list for those new and even well-established in the field. A second notable feature is the description of observed phenomena presented in the appendices (a total of 33) at the end of many of the chapters. Finally, the book is well illustrated with many figures and 88 photographs which depict the various physical phenomena described in the text.

It is unlikely that this book will serve as the main textbook for a course in cloud physics, cloud dynamics or mesoscale aspects of clouds or cloud systems. Cost alone would prohibit its general use as a class textbook. Without prior training in dynamic meteorology, radiative transfer theory and cloud physics, an undergraduate would find many of the discussions hard to follow and would be overwhelmed by the sheer volume of the material. As a graduate-level text, however, the book is limited by the lack of theoretical depth. Much of the theory in cloud physics, radiative transfer theory and cloud dynamics is that which was current during the late 1950s and mid-1960s.

To the lecturer, the book will provide excellent, easily accessible material for preparing lectures in cloud physics and cloud dynamics. Ludlam's descriptions, together with the appropriate illustrations, give form to physical concepts and provide insight into the processes governing the behaviour of clouds and related phenomena in the atmosphere.

Because this book embraces a variety of observational information not available in other texts on cloud physics or atmospheric dynamics, it will be a valuable resource book for the serious student, researcher or teacher of cloud physics or cloud dynamics. □

William R. Cotton is Associate Professor and Kevin R. Knupp a first-year PhD student in the Department of Atmospheric Science at Colorado State University.

Spectroscopy for the neophyte

SCIENCE COLLEGE
Upper Circular Road
CALCUTTA

H.A.O. Hill

An Introduction to Spectroscopy for Biochemists. Edited by S.B. Brown. Pp.403. (Academic: 1980.) £18.40, $44.50.

THE avowed intent of the editor is to provide, in a single volume, a description of spectroscopic techniques at an advanced level appropriate to the non-specialist biochemist or clinician. How closely do the individual contributors approach this estimable, if unrealistic, goal in their presentations of "non-mathematical yet rigorous" accounts of particular techniques?

The "Introduction to Spectroscopy" by the editor provides the nervous clinician with a gentle, and certainly non-mathematical, survey of the major techniques, though the physical basis of the methods could be rather better explained. The following description of ultraviolet and visible spectroscopy by Dr Brown is clear and lucid, though I would have preferred to have selection rules discussed in terms of the probability of transitions taking place. Most spectroscopic techniques are best taught by illustrating their applications, and the spectra of amino-acids, nucleosides, nucleic acids, carotenoids, flavinoids and tetrapyrroles are briefly, but adequately, presented. By far the best part of the chapter is that concerned with applications to the analysis of biological materials.

G.R. Penzer provides a near-perfect pedagogical exposition of the principles of fluorescence — the only flaw being the inadequate number of examples — and the chapter on vibrational spectroscopy by P. Gans is well-balanced with most emphasis being placed, quite correctly, on the application of Raman, especially resonance Raman, spectroscopy. P. Bayley's account of circular dichroism and optical rotation is both rigorous and non-mathematical, is well illustrated and uses many well-chosen examples, though twice or thrice as many would not have been excessive.

R. Jones continues with a description of magnetic resonance spectroscopies. These are the most difficult techniques to explain, especially the EPR spectroscopy of metalloproteins. The temptation is to adopt the simplest description even when it is wrong. Thus, though crystal field theory as an *explanation* of the consequences of introducing metal ions into a non-spherical ligand environment died 20 years ago (some would say it was stillborn), it still infiltrates otherwise respectable essays.

The book is rounded off with chapters by E.J. Wood on atomic absorption spectroscopy and by A.J. Geddes on mass spectrometry which should be valuable to those intending to apply these methods for the first time. In short, this is a book that seeks to attain an impossible goal yet with much that is good in it. □

H.A.O. Hill is a Lecturer in Inorganic Chemistry at the University of Oxford, and Fellow and Praelector in Chemistry, The Queen's College, Oxford.

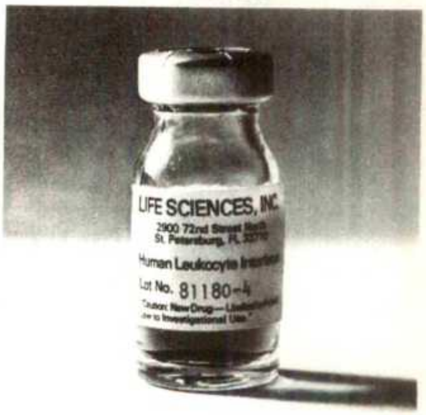
Learning from nematodes

J. Hodgkin

Nematodes as Biological Models. Edited by Bert M. Zuckerman. Vol.1, *Behavioral and Developmental Models*, pp.321; Vol. 2, *Aging and Other Model Systems*, pp.306. (Academic: 1980.) Each volume $35, £19.60.

IN STUDYING any biological phenomenon, it is desirable to work on an organism which is simple, well characterized and amenable to experimental manipulation. For many purposes, nematodes satisfy these criteria, and consequently they have attracted much attention as model systems in which to examine a variety of problems in cell and developmental biology, neurobiology and ageing. One nematode species in particular, *Caenorhabditis elegans*, has been characterized in great detail over the past 15 years, and an enormous amount is now known about this animal. However, this information is scattered widely throughout the literature, or else remains unpublished, so a monograph reviewing research on *C. elegans* and related nematodes would be very valuable. These books only partly satisfy this need, although most of the first volume, and parts of the second, are devoted to work on the genus *Caenorhabditis*. Each of the 18

⊕
LIFE SCIENCES,
INC.


LIFE SCIENCES, INC
2900 72nd Street North
St. Petersburg, FL
Human Leukocyte Interferon
Lot No. 81180-4
"Caution: New Drug— Limited by Federal Law to Investigational Use."

Supplying Human
Leukocyte Interferon for
clinical tests and
research.

2900-72nd Street North
St. Petersburg, Florida 33710
(813) 345-9371

Circle No.02 on Reader Enquiry Card.

chapters has a different author or authors, and there appears to have been little overall editorial control, so the contributions vary considerably in quality, purpose and organization. Few of them convey anything of the excitement, novelty or potential of research in this area.

The particular advantages of *C. elegans* are its neuroanatomy, development and genetics. The review of genetic analysis is excellent, and there are also adequate chapters on developmental genetics, biochemical genetics and the molecular genetics of muscle. The section dealing with cell lineages and development is long and thorough, but needlessly turgid. Neuroanatomy is scarcely described at all (apart from a good general chapter on nematode sense organs), although the total anatomy and connectivity of most parts of the nervous system have been known for some time. This omission is especially unfortunate in view of the chapter on behaviour, and the admirable summary of work on *Ascaris* neurobiology. Thus there is still no opportunity of comparing the two species without going back to the original papers.

The second volume contains contributions on ageing and longevity, nutrition, respiration, osmoregulation, energy metabolism, and the structure of nematode cuticle and sense organs. Some of these chapters are accounts of nematological research, rather than the use of nematodes to investigate fundamental biological problems, and therefore seem inappropriate to these volumes. Nevertheless, there is a great deal of interesting material in the two books; they are up-to-date and Academic Press has done a good job on the reproduction of photographs. For the time being, these volumes provide a useful introduction to contemporary research using nematodes. □

J. Hodgkin is on the staff of the MRC Laboratory of Molecular Biology, Cambridge.

Genetics and plant populations

Donald A. Levin

Demography and Evolution in Plant Populations. Edited by Otto T. Solbrig. Pp.222. (Blackwell Scientific/University of California Press: 1980.) £15, $29.50.

DURING the past decade, the discipline of plant demography crystallized from many disparate elements of empirical and theoretical ecology, and reproductive biology scattered in the basic and agricultural sciences. This discipline focuses principally on the quantitative aspects of germination, growth, reproduction, dispersal and death in a population, and the implications of these processes on the dynamics of population numbers and on life-history strategies. Although it is apparent that many characteristics of individuals related to the demography of populations are under genetic control, and that the demographic properties of populations may manifestly influence the genetic structure and amount and mobilization of genetic variation, the genetic perspective remains somewhat alien to scientists interested in plant demography. *Demography and Evolution in Plant Populations* is a step toward a unification of the ecological and evolutionary genetic components of demography.

The primary objective of the editor, Otto Solbrig, in assembling this collection of eight papers was to achieve a greater understanding of the dynamics of populations in relation to the action of natural selection. Papers by Abrahamson, Cook, Sarukhan and White deal almost exclusively with the dynamics of numbers of growth. One by Solbrig and another by Lloyd are principally genetic in content. Only the paper by Snaydon broadly integrates genetic and numerical aspects of demography. The result is a volume about demography and about evolution, rather than about the reciprocal influences of numbers and genotypes in populations.

Although the volume does not provide a broad synthesis, it does contain informative treatments of some major topics in plant demography. The high points of the volume are the treatments of numbers and size in plants by White, and the genetic variable in demography by Snaydon. Both papers bring to light much agricultural information which ecologists and evolutionists are unlikely to encounter elsewhere. Sarukhan's contribution on demography in the tropics is also packed with new information and insight. In the other chapters, the biology of seeds in the soil is imaginatively reviewed by Cook; vegetative reproduction and the genet–ramet dichotomy are treated by Abrahamson; Lloyd discusses conditions favourable to self-fertilization in predominantly outbreeding species; and Solbrig reviews the genetic structure of populations, and highlights approaches which integrate demography and genetics.

Demography and Evolution in Plant Populations will be a useful and authoritative source of data and ideas about plant demography. Scientists interested in plant population dynamics, reproduction and growth will find the volume to be of considerable interest and value. □

Donald A. Levin is Professor of Botany at the University of Texas, Austin, and specializes in plant evolution and ecological genetics.

Tanford number two

B.A. Pethica

The Hydrophobic Effect: Formation of Micelles and Biological Membranes. 2nd Edn. By Charles Tanford. Pp.233. (Wiley: 1980.) £10.50, $24.65.

THIS second edition of Dr Tanford's well-known book has been extensively rewritten to include recent advances in micelle studies, and contains a number of new chapters, most notably on serum lipoproteins and membrane proteins. It is not a review of recent results, but rather a synthesis or overview written in a lucid and occasionally magisterial style by a master of the subject.

The emphasis is on the phenomena of the hydrophobic effect, using thermodynamic arguments in a clear fashion which avoids pointless rigour. The accounts of light scattering, spectroscopic and similar methods are firm and uncluttered, and the recent rush of mathematical models receives sober treatment. The theme of the book reaches from simple solutions of amphipathic molecules to the structural basics of biomembranes, and in so slim a volume it would be difficult to please all readers with the choice of topics and facts. The omissions are almost as interesting as the inclusions. For example, bile salts are given short shrift, and the relation of dilute micellar solutions to the full phase diagrams is only touched upon. Wilkins *et al.* are quoted in support of the lipid bilayer model for erythrocytes, but Blaurock's reconsideration is not. Given that biomembranes are the sites of extensive chemical activity, the lack of reference to the now substantial body of work on micellar catalysis is a pity.

In the account of biomembranes it is good to see Robertson getting proper credit for his contribution to the bilayer membrane hypothesis. The now conventional view that biomembranes are fluid mosaics based on lipid bilayers is the one adopted by Dr Tanford, and to me the important last chapter (on membrane proteins) consequently has the air of skilful fitting of the hydrophobic proteins to a Procrustean bed. Perhaps when physicochemical data on the proteins are as extensive as for the lipids, the remarks of J.B.S. Haldane (quoted at the front of the book), which identify the membrane polymers as the key components, will be seen as truly prescient and the role of the lipids will be put into better perspective. Be that as it may, the hydrophobic effect will remain of major importance in the study of lipids and proteins alike, and this book will beneficially influence future research on membranes and other structures. We should hope for a third edition for the pleasure of reading Dr Tanford again. □

B.A. Pethica is at Clarkson College, Potsdam, New York.

nature

12 February 1981

Who now believes in the neutron bomb?

The Reagan Administration, which has been promising clarity and decisiveness since it occupied Washington three weeks ago, has begun badly in its pronouncements on the neutron bomb. One day, there was Mr Caspar Weinberger, the Secretary of Defense, promising that plans to deploy what is called the neutron bomb in Europe would be dusted off and reconsidered. A few days later, last week, Mr Alexander Haig, conscious because of his long spell in Brussels of European susceptibilities on this point, was saying that no decision had yet been taken and that none would be without prior consultation with the European members of the North Atlantic Treaty Organization. People in Europe will recognize that they have been here before. It is only two years since President Carter twisted every arm in sight, persuaded most but not all members of the alliance to accept the notion that neutron bombs were a sensible part of the armoury of the West — and then backed down at a point at which the West German government was in the thick of persuading a sceptical Bundestag that Mr Carter's first thoughts were right. Mr Haig is right to have spotted so quickly that the same battle cannot be fought a second time without spilling somebody's blood; perhaps he will make a good Secretary of State. Mr Weinberger's ruminations are understandable — a new man in an old post, discredited by too much vacillation, anxious to give the new Caesar comfort. They are not so easily forgivable.

The issue of the neutron bomb is more than merely a problem in military technology, but the technology is not irrelevant. Neutron bombs, when first described in outline three years ago, were held to have special military advantages. They could, the argument went, kill the crews of tanks whatever the thickness of the armour protecting them from the effects of conventional explosives. In principle, the argument is credible. The assumption must be that it is possible to generate neutrons from fusion explosions (the energy of fission neutrons is too little) without generating too much of the encumbering mechanical and thermal energy. Neutrons between ten and twenty million electron volts constitute a penetrating form of nuclear radiation. There is a good chance that an impinging neutron will penetrate the few inches of armour plating around a tank. The people inside, however, are more vulnerable. Consisting (as they do) largely of water, people function much like the moderating materials in nuclear reactors. They decelerate neutrons more efficiently than armour plating, chiefly because they have lighter atomic nuclei, but only by generating within themselves fast-moving and potentially damaging protons. The result is that people who think themselves protected by the vehicles they ride in will often be exposed to a fatal dose of radiation. In the macabre nature of these happenings, however, they will not always be aware of what has hit them. The first nausea may not come over them for half an hour or so and by then, of course, they may be dead from other causes as people on battlefields often discover to their cost.

Technically, the neutron bomb is therefore an exceedingly specialized weapon. It cannot be used on a battlefield as such, for its effects would be too slow. It is not, however, a strategic weapon. Yet the effects of a neutron bomb must be largely local. A hundred metres of atmosphere would probably have a stopping power for the spectrum of neutrons from a thermonuclear mixture not very different from a thin skin of armour plating. Thus pure neutron bombs cannot be used to attack targets spanning more than a few hundred metres on the ground without losing much of their power to do damage. On the other hand, they cannot be detonated near the ground without generating fallout. This, no doubt, is why most speculation on the subject has been concerned with the use of neutron bombs against concentrations of tanks marshalled for conventional attacks on some nearby target. By that means, the argument goes, it might be possible to prevent the tanks from rolling without so polluting the neighbourhood that everybody would be hurt. The scenario is appealing but not persuasive. Will generals saddled with the responsibility for using neutron bombs be able to make the delicate calculations of optimum detonation height that nuclear physics will require of them or their assistants? Will their opposite numbers in command of the hypothetical hostile tanks, having read in the newspapers of neutron bombs, then calmly assemble their forces in the manner appropriate to the Second World War? And, given the difficulty of arranging for public demonstrations, will either side in some future conflict take these drawing-board weapons seriously?

What Mr Haig to his credit appreciates is that there are more immediate problems. With the passage of time, it becomes more apparent that each new weapon requires a separate negotiation within the alliance which is the chief permanent commitment of the United States outside its borders. The wish to site cruise missiles in European states is still being talked about, and in some places (the Netherlands) may yet be talked out. The case for siting neutron bombs (Carter-style) in Europe took almost as long to make before President Carter listened more attentively than he need have done to what Mr Brezhnev had to say on the subject. In reality, the United States need not have made an issue of the neutron bomb and, more to the point, should not have done so. Cruise missiles raise novel issues, but neutron bombs are covered by the agreement within the North Atlantic Treaty Organization on the deployment of tactical nuclear weapons. If, as seems likely, they have only a specialized role to play in defence planning, and if the consequences of their use are likely to be less and not more serious, no new principle seems to be involved. In short, the issue of the neutron bomb seems less a consequence of improved technology in nuclear weapons than of President Carter's wish to say something publicly that had not been said before. Mr Weinberger appears to have fallen into the same trap.

Mr Haig's proposed remedy makes the best of a bad job. He promises that there will not now be a decision to site neutron bombs in Europe without full consultation with Western European governments. Given the history of the past few years, he has no other diplomatic choice. Mr Brezhnev should be pleased, but will refrain from saying so, even if he knows that his complaints at President Carter for seeking to deploy neutron bombs were partly bluff. (Neutron cross-sections are the same in the Soviet Union and the United States.) Mr Haig will no doubt also be pleased; he will have etablished his authority over the Department of Defense early in the new administration. Yet the problem of Europe remains to haunt him. Only three months after being returned to power, the West German coalition government is now less able to face a political battle about novel kinds of weapons than it was before the election last November. Other members of the alliance, the Netherlands for example, appear to have given up the internal struggle. None of this implies that NATO is about to fall apart but only that many of its members have become aware of the reality of the risk of conflict, and of the upheavals that may yet follow what is happening in Poland. Mr Haig must want to bring Western Europe up to scratch. Paradoxically, he may find that the only way to do that is to dust off not the neutron bomb but the series of arms control negotiations between East and West that have been in limbo since last September.

0028-0836/81/070523-02$01.00

New Reagan budget a mixed blessing

Social science and planetary missions cut

Washington

President Ronald Reagan's first budget proposals are likely to contain a mixed message for science when they are formally unveiled next Wednesday. The emphasis is likely to remain on long-range research of potential value to the economy, and basic research of less direct relevance — such as planetary exploration and the social sciences — seems likely to suffer hefty cuts.

These are the main themes of preliminary proposals being circulated by Mr David Stockman, director of the Office of Management and Budget, as part of an effort to boost the US economy by large-scale reductions in social spending.

As expected, most of these cuts will fall on social welfare programmes, for example reducing health benefits and support for education. This theme will be reflected in the budget of the National Science Foundation (NSF) where major cutbacks are being proposed not only in research funding in the social and economic sciences but also in science education and support for women and other minority scientists.

Whereas President Carter, for example, shortly before leaving office had proposed a 1982 budget for NSF that would have allocated $84 million for research financed through the Behavioral, Social and Economic Sciences Directorate — an increase of about 15 per cent over the 1981 figures — Mr Stockman has proposed that funding for social and economic sciences be reduced from $33.6 to $10 million, and for behavioural and neural sciences from $39.6 to $30 million. Similarly he is proposing to eliminate the Women in Science programme, introduced last year.

In contrast to these cuts, it is intended that former President Carter's proposals for increased spending in the so-called hard sciences, such as physics and chemistry, should remain intact. These increases had been justified as part of the Carter Administration's attempts to accelerate the pace of technological innovation. And since this is likely to be a top priority for the Reagan Administration, it could therefore become the chief function for NSF, which has in recent years seen different types of goals added to its original mandate.

So although Mr Stockman is asking for an overall cut of $240 million in NSF's budget from that suggested by Mr Carter, the proposed increases for the hard sciences would remain untouched, leaving the agency with a slight growth over the year.

It is likely to be different at the National Aeronautics and Space Administration (NASA), where budget office is proposing virtually to eliminate the planetary science programme in its efforts to cut the agency's budget while keeping the space shuttle — now scheduled for an April launch — on target.

In particular, Mr Stockman is proposing to eliminate the Galileo mission to Jupiter — already running into rescheduling problems because of the difficulties in the development of the inertial upper stage needed for launching from the space shuttle, and the resulting decision to use a Centaur launcher.

Budget cuts of this order would result in a long period in which the United States was not engaged in planetary exploration. They are likely therefore to be bitterly contested by the scientific community and its supporters in Congress.

The proposed cuts in NSF's programmes are also expected to meet fierce opposition. When President Richard Nixon tried to inflict similar cuts, his efforts were successfully resisted by Congress, although the then Democratic Senate was a much greater hurdle than its present Republican counterpart.

No nominations have yet been made by the Reagan Administration for top science posts, although the President's team would probably like to have some individuals in position before the budget cuts begin.

Favourite to become director of the Office of Science and Technology Policy — which Mr Reagan is said to have been persuaded to retain — is Dr Simon Ramo, co-founder of TRW Inc. and head of Mr Reagan's Science and Technology Task Force set up immediately after last November's election. If Dr Ramo rejects the offer said to have been made to him, the choice could be Dr Arthur Bueche, vice-president for research of General Electric.

At NASA, the present favourite for administrator is Dr Hans Mark, former head of NASA's Ames Research Center and Secretary of the Air Force under Mr Carter. Dr Mark is said to have impressed Mr Reagan's transition team, and would be a popular appointment at the agency.

David Dickson

European high-energy physics cuts

European high-energy physics is heading for trouble. The Deutsches Elektronen Synchrotron laboratory (DESY) in Hamburg faces the prospect of operating for only 4–5 months a year, and the European nuclear research laboratory (CERN) at Geneva is likely to find its plans for a new electron–positron machine postponed for at least 6 months.

In Germany, the research ministry appears to have decided that, after some years of growth, the DESY budget can be cut back. And in Sweden, the Ministry of Research and Education is "suspicious" of the haste at CERN to build the 30-km electron–positron storage rings, with an initial electron-beam energy of 50 GeV. One source of trouble is that many Swedish physicists at CERN work not on high-energy physics but on spin-off experiments with the low-energy accelerators that litter the CERN site. Although not glamorous, such experiments are extremely sophisticated, and in danger of being neglected.

The Swedish Ministry of Research and Education said this week that it was seeking information from CERN on the future of these activities. Acknowledging that Professor Herwig Schopper, director-general of CERN, would like a decision on the large electron–positron machine (LEP) at the next bi-annual CERN council meeting in June, a Swedish spokesman asked "Why are we in such a hurry?" The intention at CERN is that LEP would be built within the running budget of 610 million Swiss francs a year, but Schopper freely admits that this would entail running down some ancillary facilities.

Although the annual running cost would not be changed, Stockholm is alarmed that the usefulness of CERN to Swedish scientists will diminish. In any case, a commitment for 15 years or more to the construction and running of LEP would have to be ratified by the Swedish parliament, said the spokesman, and that was most unlikely before June.

Schopper's dilemma is to reconcile the interests of governments broadly willing to support LEP but at a smaller budget (SF590 million a year) such as France, Germany and the United Kingdom and those of the smaller nations wishing to keep less costly experiments going. one sop might be to squeeze another SF95 million or so from the budget to build the European Synchrotron Radiation Source, the machine proposed by a panel of the European Science Foundation. "That would be nice" said the Swedish spokesman.

At DESY, ranking second only to CERN in Europe, the federal government is proposing to slash the budget for 1981 to DM135 million, compared with DM145.4 million last year. The government has met DESY's salary bill but made no allowance for increases (5 per cent would be necessary for inflation); cut DESY's budget for consumables by 10 per cent in money terms, making no allowance for the doubling of Hamburg electricity prices last September; and cut DESY's provision for investment in detectors and accelerator improvement by 22 per cent.

The laboratory can cope with the

0028-0836/81/070524-02$01.00

electricity prices (caused by the Polish crisis, which has reduced Hamburg's supply of cheap coal to a trickle) by reducing its operating time to 4½ months (in 1980 it was 9 months). At present the laboratory is counting on a partial reprieve and working as if it will get 6 months. If it does not — the Bundestag will decide in the spring — it will be quiet in Hamburg later in the year. **Robert Walgate**

Soviet agriculture

Soil first

Dr Vladimir Borovskii, director of the Kazakh Institute of Soil Science, has strongly criticized the All-Union Academy of Agricultural Sciences for its neglect of soil problems. The academy, he said, has no long-term programmes for studying the nature of the soil and establishing the laws of soil formation.

Interviewed on Moscow radio last month, Dr Borovskii said that unless these shortcomings are put right, the development of Soviet agriculture will inevitably be retarded. The draft basic guidelines for the new Five Year Plan, which are to be presented to the Party Congress later this month, should, he urged, include a clause that soil research must be considerably expanded.

The weeks between the publication of the draft guidelines and the meeting of the Party Congress which will approve them is traditionally a time for constructive criticism. As always, the guidelines emphasized the needs of agriculture.

By 1985, it is planned that production of mineral fertilizers should reach at least 150 million tonnes annually, at least 115 million tonnes of which will go to Soviet agriculture. A further 3.4–3.6 million hectares of irrigated land and 3.7–3.9 million hectares of drained land are to be brought into use. Some 26–28 million hectares of pasture are to be reclaimed from desert, semi-arid or mountainous regions. There is to be a drive for more efficient use of irrigated or drained land, and improvement of the technical level of water conservancy. Improvements are also expected in reducing the salinity of agricultural water supplies and soil acidity.

Mr Brezhnev clearly had these measures in mind in his message to the All-Union Agronomists' Conference in Moscow last December. Soviet agriculture, he said, had "powerful levers for raising productivity — modern technology, chemical means, large areas of irrigated and drained land, the achievements of science and advanced experience". In a similar vein, Politburo Member Mikhail S. Gorbachev stressed the "profound learning, highly scientific spirit and professional competence of Soviet agronomists". It was when the advice of such experts was ignored, he told the conference, that things began to go wrong.

Gorbachev's main criticisms of Soviet agriculture related mainly to planning. What was needed, he said, was "a sharper turn of the whole economic, organizational and political work towards effective use of the material, labour and financial resources", the more rational use of land, substantial increases in the production of grain and fodder and "solution of the social problems of the countryside". **Vera Rich**

Solar power

Crimean costs

The Soviet Union's first solar power station will go into operation in 1985, according to a "round table" discussion on solar energy published last month in the Moscow weekly *Literaturnaya Gazeta*. This pilot station, to be built in the Crimea, will have a generating capacity of 5 MW, and will run at an estimated cost of 1.5–2 rubles per kWh.

The cost is high. According to the Soviet Ministry of Energy, in 1979 the production cost of 1 kWh was 0.752 kopeks at a conventional thermal station, 0.786 kopeks for nuclear power and 0.149 kopeks for hydroelectricity. Thus the generation cost of Crimean solar electricity will be almost two hundred times that of nuclear power and more than a thousand times that of hydroelectricity.

Nevertheless, the pilot station has an important place in the new Five-Year Plan, which calls for an increase of renewable energy use in the national economy. While maintaining that the Soviet Union runs no risk of a shortage of conventional fossil fuels, energy planners have for several years been insisting on the need to conserve them, stressing their irreplaceable role as feedstocks for industry.

A number of solar energy projects are already under way in the Crimea, which has an average of only 47 days a year without sunshine. These include the Alushta experimental centre, where a village of single-storey cottages is used to test solar-powered domestic heating/air-conditioning units, and an experimental metallurgy centre at Karsiveli where parabolic mirrors will be used to produce temperatures of 3,500°C, allowing metals to be smelted to a high degree of purity.

The intended solar power station will also use parabolic mirrors, heating water to drive steam generators. Long-term plans for solar energy will take more sophisticated forms. According to Academician Nikolai N. Semenov, academic secretary of the "scientific council for finding new ways of using solar energy", research is being concentrated on photoelectric conversion.

Chemical methods of dissociating water are also being explored, according to Semenov, whose contribution to the round table differed very little from what energy strategists elsewhere are accustomed to say. Semenov emphasized the need for continuing research at a basic level in the exploitation of solar energy. **Vera Rich**

Wildlife conservation

Keepers' charter

The new British Wildlife and Countryside Bill has accomplished one thing at least — it has attracted a record of 600 amendments for debate in the committee stage in the House of Lords. After two weeks, the number of amendments discussed by the end of last week was well below target. The government still has a long haul ahead.

Most of the amendments dealt with so far have been withdrawn for further consideration by the government, whose views will not be known until the bill emerges from the committee. The government has however agreed even at this early stage to include the red squirrel, now down to a handful of breeding colonies in England, on the list of protected species. Other amendments, which have aroused varying degrees of interest but as yet no firm government response, include protecting bats in private houses, banning the use of crossbows to kill wild animals and protecting badgers and red foxes from the hazards of snares.

The debate is likely to become more fierce later this week, when the House of Lords turns from debating the protection of wild animals and plants (the first part of the bill) to the conservation of the countryside and national parks. The bill includes a clause relating to sites of special scientific interest, designated as such by the Nature Conservancy Council. Voluntary conservation bodies and the council have condemned the draft clause which requires that the owners of only some specially selected sites will have a statutory obligation to give notice of operations that could alter the characteristics of the site.

They argue that all sites should have the same protection and that the singling out of some sites would lower the status of the others. The clause has provoked a spate of amendments, the more extreme of which ask that all agricultural practice should be subject to the planning laws. The strength of opinion is likely to persuade the government to cave in.

Many of the amendments now being tabled could change the status of the Nature Conservancy Council, statutorily responsible for conservation. Not only might the council be given more power in enforcing protection for sites of special scientific interest but, if the voluntary bodies have their way, it could be called upon to help enforce wildlife legislation generally.

The council's reaction to the voluntary bodies' enthusiasm for awarding it greater power is cautious — it says simply that it is primarily a scientific organization. Memories of the bruising battle with the Natural Environment Research Council a decade ago, from which the conservancy council emerged the loser, are still fresh. **Judy Redfearn**

0028-0836/81/070525-02$01.00

German nuclear power

Local difficulties

Hamburg

West German Chancellor Helmut Schmidt will need to do some political juggling in the next few weeks if he is to maintain the credibility of plans to expand West Germany's nuclear energy programme to lessen dependence on oil imports. The dilemma stems from a decision last week by local members of his own Social Democrat Party (SPD) in Hamburg that the city should not contribute 50 per cent of the costs of a new nuclear power station planned for Brokdorf, 40 miles away.

Despite support for the power station from several local and national trades union, the Hamburg SPD accepted the arguments of the city's mayor, Herr Hans Ulrich Klose, that it was more important to base future energy policy on energy conservation than to support significantly increased production of electricity by nuclear means.

The Brokdorf plant, which would be built and operated by the Hamburg Electrical Utility, has become something of a symbol for the German anti-nuclear movement. Two years ago it was the scene of the first of a series of violent demonstrations against nuclear power which culminated in the decision to withdraw plans for a major storage and reprocessing facility at Gorleben.

Chancellor Schmidt, in putting considerable political weight behind the Brokdorf proposal, seems to have been gambling on an apparent decline in popular anti-nuclear sentiment over the past two years. The link between growing economic problems and OPEC oil prices is increasing pressures to proceed with the construction of nine new nuclear power stations.

Mayor Klose's arguments against Brokdorf were expressed more in local than national terms. He pointed out, for example, that although the country as a whole at present relies on nuclear energy for only 10 per cent of its energy needs, the completion of the Brokdorf plant would increase Hamburg's dependence from 30 to 70 per cent.

"How can we encourage energy saving when we blindly push ahead with nuclear power that may well prove to be superfluous to our needs?" Herr Klose asked a meeting of the local SPD, in a debate which ended with a vote against the plans.

The decision is not yet final, since it will have to be taken formally by the Hamburg Senate, which meets this week to vote on the city's participation in Brokdorf. However, it is unlikely that the view of the ruling SPD will not prevail.

The Brokdorf situation has already become a major political embarrassment to Chancellor Schmidt, who has made a commitment to other European leaders that his nation will make a major effort to reduce its dependence on foreign oil.

The SPD's position is that, pending the resolution of the waste disposal issues, the construction of nuclear power plants already under way should continue, as long as an adequate form of temporary disposal has been agreed.

In Hamburg, a plan for dealing with waste has been produced that seems to meet the conditions laid down by local port authorities. But the local SPD's decision — which runs directly contrary to Herr Schmidt's plans — has already been described by Hamburg's Interior Minister as a possible precedent for other nuclear stations.

The political problems raised by Hamburg's actions have further exacerbated tensions at a national level, where Herr Schmidt needs a viable energy policy to maintain the SPD's coalition with the Free Democratic party, but is finding his position on nuclear energy out of line with that of a large proportion of his own party.

For example, in Hamburg's neighbouring state Schleswig Holstein, the national SPD feels it has a chance of defeating the ruling Christian Democrats in the next elections. However, as in Hamburg, the local SPD is also opposed to the Brokdorf plans.

Not surprisingly, the national SPD has been somewhat embarrassed by Herr Klose's firm anti-nuclear stance. Some local SPD members who do not support the mayor's position had been hoping that he would be replaced by Herr Hans Apel, currently Defence Minister in Bonn, who supports Chancellor Schmidt's position. However, other political pressures make it unlikely that Herr Schmidt would agree to such a weakening of his own cabinet.

David Dickson

EEC nuclear research

Reactor guarded

Brussels

The go-ahead on the European Community's light water reactor (LWR) experiment, the Super-Sara project, is expected to be given at a meeting in Brussels this week. The ten member states' permanent representatives to the Community (Coreper) will be deliberating on a report by the Council of Ministers' Working Group on Atomic Questions before deciding whether or not to release the main bulk of the funds for the three-year programme (1981–83).

In March last year, the Nine, on the insistence of the French, allocated an initial grant of 3.31 European Units of Account (EUA), or £1.7 million, for the exploratory stage of the project. During the previous six months, the member states had been deeply divided as to whether the research was really necessary or justified at the price. Initial estimates put the cost of the programme at 43.92 million EUA but inflation and salary increases have now increased this to 54.03 million EUA.

The project makes use of the Community's Essor experimental reactor at Ispra in Italy, which, having served its original purpose, would otherwise become redundant. Italy is unwilling to take over financial responsibility for Essor and through the Super-Sara project, drawn up in 1979 in the wake of the Three Mile Island accident, the Commission is willing to take this on. It aims to simulate as far as possible the complete gamut of accidents which might occur in light water reactors. More extensive than similar projects in West Germany (the Rebeka reactor), France (Phébus) and the United States (PBF and MRBT), the simulated accidents include loss of coolants from small and large cracks, the exposure and fusion of the core, the effect of the containment breaking up, the monitoring of events in and outside the core and the possibilities of recooling the core.

The real task of the Commission since the project was agreed in principle last March has been to define the areas where the Community could benefit. The report which formed the basis of the Commission's communication to the Group on Atomic Questions was drawn up by a task force composed of experts from the member states, and the decision taken at the Coreper meeting will reflect the extent to which the task force has been successful.

The United Kingdom's original position — that the project is a good idea but in practice would be too expensive — changed after the March meeting to one of acceptance provided that the project went ahead as soon as possible. At that time, other member states, Germany and the Netherlands in particular, proposed carrying out the research in two stages and withholding much of the money until Super-Sara had proved its value. But postponing the project for two years, argues the Commission, would cost an extra 32 million EUA. Moreover, a follow-up programme planned for 1983–86 might produce particularly valuable results. France and many other countries are more worried that the project's costs will rise more steeply than predicted, with the high temperatures at which the experiments take place. France also feels that its own experimental reactor, Phébus, although smaller, will produce similar results just as quickly. On the other hand, the United Kingdom may now be prepared to come to the programme's defence in view of the recent decision to build the first British pressurized water reactor in the face of strong public opposition, and because of the United Kingdom's lack of safety expertise with light water reactors.

The Commission has been at pains to emphasize that it will ensure "a particularly rigorous control" of the costs and timetable. Another possible line of attack for the programme's detractors is the lack of qualified staff. An estimated 70 temporary staff will be needed in addition to the 2,600 permanent staff. However, the Commission emphasizes that other nuclear safety programmes will not be robbed of their present staff.

The biggest arguments in Coreper are likely to revolve around the timing of the project. Work on the loop has already begun at Ispra and to keep Essor unoccupied now will only lose money. But while the Commission is urging that the project should start in earnest immediately, the nuclear safety programmes' consultative committee wants the project's first phase to be extended for 6–12 months so that more feasibility studies can be carried out. However, financial support from third parties may tip the balance in favour of an early start. Three United States organizations — the Department of Energy, the Nuclear Regulatory Commission and the Electric Power Research Institute — are all considering participation, as are Japan and Sweden.

Jasper Becker

Data protection laws

UK lags behind

The United Kingdom seems on the point of becoming a "data desert" following its decision not to sign the Council of Europe's convention on data protection and privacy. On 28 January, seven European states signed the convention, which requires them to enact legislation defining the rights of individuals over the formation and use of files concerning them. All these states except one (Turkey) already have such laws. When ratified by the parliaments of five signatories, the convention will pass into force.

Britain must establish a domestic policy first, said ministers questioned in Parliament last week. But the issue has been under debate in Britain for ten years, and it is two years since the Lindop committee on data protection produced its report and recommendations. Mr William Whitelaw, Secretary of State for the Home Office, last October promised a policy statement "shortly". But there is now little sign that a statement will be forthcoming in the present session of Parliament.

It is a matter of human rights, but also of commercial interest. Companies argue that contracts involving protected data from outside countries will be lost to the United Kingdom, as Britain will be seen as a potential information leak. And this information can concern companies as well as individuals: the Council of Europe convention is defined in terms of "legal persons", which may include registered businesses.

Already, according to International Computers Ltd, a number of small deals with Sweden — the first in Europe to enact data laws — have been affected. Austria is also beginning to pay close attention to the nature of the data that cross its borders. So far there has been no serious commercial effect; but the European laws are only just beginning to operate.

In the United Kingdom, bodies from IBM to the National Council for Civil Liberties (NCCL) support the spirit of the Lindop report — that there should be legislation within five (now three) years. Opposition stems from government departments whose data handling might be most affected by legislation: the Home Office and the Department of Health and Social Security. The Department of Industry backs the data industry's view that legislation is necessary, although the appointment last year of a junior minister for information industries (Mr Adam Butler) specifically excluded data protection from his portfolio.

Lindop recommended the establishment of an agency which would initially coordinate views and propose legislation to Parliament (thus by-passing departments and ministers). Later the body would register and license protected files and uses of files.

Such an authority is regarded by NCCL as a necessary democratic safeguard. But commercial interests, while recognizing the need for independence, are fearful of a centralized bureaucracy. IBM, involved in most countries, favours the decentralized German system. There, firms and government agencies holding sensitive files must appoint an ombudsman who then bears legal responsibility for ensuring that files are properly constituted and used. Costs are thereby distributed to the file owners, rather than being a matter for central government.

The system also avoids the danger — so IBM thinking goes — of a "Balkanization" of data transfers, which might occur when national protection agencies attempted to agree on this or that data flow between their countries. Another way out, however, would be the wide adoption of a convention such as that drawn up by the Council of Europe.

Broadly, the Council of Europe convention — which is open to states outside the European region — commits ratifying states to enact legislation requiring that personal data be lawfully obtained, accurate and relevant, and that its use be restricted to approved purposes. Any information recorded on race, politics, religion, health, sexual behaviour and criminality is subject to extra safeguards — although the convention may be over-ridden in extreme cases involving state security, public fraud and crime.

The signatories so far are Austria, Denmark, France, Germany, Luxembourg, Sweden and Turkey. Within Europe, Norway also has data protection legislation but has not yet signed the convention. Outside Europe, the United States and Canada have legislation. It is being considered at one level or another in most developed states. **Robert Walgate**

European astronomy

Winning ways

Hard on the heels of the race for the Space Telescope Institute, won by Johns Hopkins University last month, comes a similar competition in Europe. Four astrophysical institutes are competing for the European Coordinating Facility that will analyse and store space telescope data for European astronomers. The winner is expected to be announced by the European Space Agency (ESA) in June.

The competitors are the Royal Observatory, Edinburgh, the European Southern Observatory, Munich, the Observatoire de Paris and the Italian Institute for Astrophysics, Frascati. They have submitted bids to have thier own hardware used to store the data and develop programs for analysing them.

European astronomers will be allocated at least 15 per cent of the space telescope's observing time. Observers, who will only be a fraction of those wishing to use the telescope, will typically spend one or two weeks at Johns Hopkins and then return to Europe with their data for further analysis. Astronomers working at institutes with hardware compatible with that of the coordinating facility will be able to use its programs either by accessing them directly through computer links or by applying for tapes through the post.

Observers will have exclusive rights to their data for one year after which they will be available to the whole astrophysical community. The National Aeronautics and Space Administration and ESA have recently agreed that all space telescope data should be stored at the European facility as well as the Space Telescope Institute.

The European facility will most probably use a VAX 11/780 computer manufactured by the DS Digital Corporation. In an effort to increase the compatibility of hardware, several astrophysical institutes in Europe have decided to adopt this computer over the past year. Of the four competitors for the facility, the Royal Observatory, Edinburgh, with links to five other nodes in Britain, has the most highly developed system based on the VAX 11/780, making it perhaps the most likely winner. But the other competitors could catch up. There are plans to link Frascati's computer to other nodes in Italy and several computers compatible with VAX 11/780 are already linked in France.

Most of the cost of the coordinating facility will be met by the chosen institute. Although all candidates have suitable hardware, some could be required to expand their facilities more than others, involving them in considerable expense. ESA will provide only modest support: about £50,000 towards start up costs and thereafter £25,000 a year for incidental expenses plus seven staff salaries.

Judy Redfearn

0028-0836/81/070527-01$01.00

CORRESPONDENCE

Selfish genes in race or politics

SIR — I hesitate before giving the National Front even a tiny bit more exposure in your pages: I imagine they welcome verbal denunciations, much as, on a larger and a physical scale, they presumably relish a "Smash the Front" march. But a public challenge from Steven Rose has just that touch of insinuation which makes me feel I would be unwise not to respond. He says (*Nature* 22 January, p.335): "May I suggest that it would be in the public interest that John Maynard Smith and Richard Dawkins should clearly dissociate themselves from the use of their names in support of this neo-Nazi balderdash".

Balderdash is a well chosen word. The extract from *New Nation* that he quotes is of course obvious balderdash, a travesty of science, in the service of an evil cause. Rose would have had every right to have added, "I told you so." He and his colleagues have long urged that "in due course racists would deploy sociobiology in support of their views". In my naive way, I never believed it possible that such a deployment could seem logical to even the most twisted mind, but perhaps I just lacked Rose's intimate knowledge of how science can be misused for political ends.

The equating of "kinship", in the sense of kin selection, with "ties of race" appears to result from an interesting variant of what I have called the fifth misunderstanding of kin selection[1].

Let me pass from this relatively esoteric point to concentrate on the central fallacy of the passage from *New Nation* that Rose quotes: ". . .our genes don't permit us to live in a Marxist–Rousseauesque egalitarian communist utopian World State of universal altruism. It was an inevitable result of the way evolution works that our genes would not permit us so to live."

The key to the error is to be found in the word "inevitable". I tried to make it clear in *The Selfish Gene* that ". . . it is a fallacy — incidentally a very common one — to suppose that genetically inherited traits are by definition fixed and unmodifiable. Our genes may instruct us to be selfish, but we are not necessarily compelled to obey them all our lives. It may just be more difficult to learn altruism than it would be if we were genetically programmed to be altruistic".

Insofar as my book has any political moral, it is conveyed by the closing sentence: "We, alone on earth, can rebel against the tyranny of the selfish replicators".

There is no evidence that racism is a naturally selected tendency, but just suppose some such evidence were to be discovered. What should be our response to it? The *New Nation* author presumably would say: Racism is the product of genetic evolution, therefore it is inevitable and desirable, and should be formalized in our political institutions. I would say: Racism is highly undesirable, and if it is indeed the product of genetic evolution this simply means that we should fight all the harder against it. There is a third position, untenable but, nevertheless, commonly held: Racism is highly undesirable, therefore we cannot allow it to be true that it is the product of genetic evolution. There must be something wrong with the evidence.

What is really wrong with the National Front quotation is not the suggestion that natural selection favoured the evolution of a tendency to be selfish and even racist. What I object to is the suggestion that if such tendencies had evolved they would be inevitable and ineradicable: the suggestion that we are stuck with our biological nature and can't change it. It is this that is the real balderdash, but where on earth did the myth of the inevitability of genetic effects come from? Is it just a layman's fallacy, or are there influential professional biologists putting it about?

Consider the following passage from Rose's[2] review of E.O. Wilson's *On Human Nature*: ". . . although he does not go as far as Richard Dawkins . . . in proposing sex-linked genes for "philandering", for Wilson human males have a genetic tendency towards polygyny, females towards constancy (don't blame your mates for sleeping around, ladies; it's not their fault they are genetically programmed)".

We may pass quickly over the somewhat free interpretation of my book. "Philanderer males" were not humans, they were hypothetical animals postulated as part of an elementary mathematical model, and in any case the computer simulation which I described ended up with the philanderers slightly outnumbered by "faithful" males.

The serious point I want to make arises out of Rose's amusing remark about ladies not blaming their mates. Just suppose it were shown that "human males have a genetic tendency towards polygyny, females towards constancy", what on earth would that have to do with concepts like "blame" and "fault"? Everybody knows that most healthy humans of both sexes have strong sexual desires which are not always strictly monogamous in aim, and many humans nevertheless enter into voluntary contracts of sexual fidelity, with some at least momentary intention of overcoming their polygamous tendencies. Many even succeed in this.

Depending on our philosophical position with respect to "free will", on the nature of the personal relationships concerned, and on the form of the contract entered into, words like "blame" and "fault" may or may not seem appropriate. But whether or not they are appropriate will not be altered one whit by anything that geneticists may ever discover. If sexual desire is strong or weak, it is so whether or not we label it genetically determined.

The same goes for racial prejudice. Whether it is eradicable or not is a serious question, but the consideration of whether it is "genetic" has no bearing on the matter. Rose may have forgotten the difference between "sex-linked" and "sex-limited" (see quotation above), but he will certainly recall that warhorse of the genetics textbooks, phenylketonuria.

This serious disease, caused by a single recessive gene, is easily cured by rearing the child on a special diet: an "environmental" cure for a genetic disease. If racism should turn out to be a genetically determined disease, I see no reason why it, too, should not be easily cured by suitable rearing conditions. And if it is not easily cured by educational means, exactly the same might be true of an acquired prejudice, say a prejudice that one has been taught during some critical years of early life. There is nothing particularly irrevocable about genetic effects on development, as compared with non-genetic effects.

How Steven Rose and the spokesmen of the National Front came to share their fatalistic views on the inevitability of genetic determination I cannot guess, unless it has something to do with the fact that "historical inevitability" is as dear to the Marxist heart as the related concept of "destiny" is to the Nazi one. In their biological manifestations, at least, the two concepts are as fatuous as each other.

But let me sound another note that may strike a chord in Rose's heart, a note of pure political expediency. If a scientific theory, X, about the nature of man, could have evil human consequences, we may hope that X will turn out to be false, but we would be unwise to take our stand purely on our hope that it is false. After all, it may just turn out to be true, and then we are left helpless and with our guard down. It is much wiser to say: "I don't think X is scientifically valid, but even if it *is*, so what?"

This message is nowhere more important than in our dealings with the other controversy mentioned in Rose's letter, the race/IQ controversy. As Stephen Jay Gould[3] puts it with his customary cogency: "I do not claim that intelligence, however defined, has no genetic basis — I regard it as trivially true, uninteresting, and unimportant that it does . . . It is just as likely that blacks have a genetic advantage over whites. And, either way, it doesn't matter a damn".

In the aftermath of Mrs Thatcher's electoral victory in 1979 Steven Rose wrote[4]:" . . . the switch in scientific fashion, if only from group to kin selection models in evolutionary theory, will come to be seen as part of the tide which has rolled the Thatcherites and their concept of a fixed, nineteenth century competitive and xenophobic human nature into power."

The theory of kin selection is logically entailed by the now virtually undisputed neo-Darwinian theory itself. It explains things that have been going on in the world for a thousand million years, and that will go on after our species is long extinct. Versions of it probably hold sway in uncountable islands of life all over the universe. It is annoying to find this elegant and important theory being dragged down to the ephemeral level of human politics, and parochial British politics at that. It seems that the National Front are not alone responsible for this.

RICHARD DAWKINS

New College,
Oxford, UK

---

1. Dawkins, R. *Zeitschrift Tierpsychol.*, **51**, 194-200.
2. Rose, S. *New Scientist*, 5 October, 1979.
3. Gould, S. J. *Ever Since Darwin*. Burnett, London.
4. Rose, S. *New Scientist*, 17 May, 1979.

NEWS AND VIEWS

Betty 15/2/82

A new measure of intelligence?

from N.J. Mackintosh

RECENT experiments by Christopher Brand at the University of Edinburgh have revealed an extremely high correlation between IQ and performance on a very simple task — that of judging which of two briefly presented lines is the longer. Subjects seated in front of a tachistoscope are shown, for example, two vertical lines side by side, one approximately 50 per cent longer than the other. Their task is simply to say whether the longer line is on the left or the right. If the lines are displayed for more than a few tenths of a second, the task is, of course, trivially easy. But errors can be induced by reducing the duration of exposure. 'Inspection time' can then be defined as the minimum exposure duration at which the subject can maintain a level of accuracy of, say, 95 per cent. A series of studies by Brand[1] and earlier work by Nettlebeck and Lally[2] at Adelaide show remarkable correlations, ranging up to -.92, and with an average of about -.80, between inspection time and various measures of IQ.

These correlations demand to be taken seriously. According to Brand, they suggest the possibility that intelligence is really a matter of quickness of apprehension. The quick child becomes the intelligent adult because he has been able to learn more, and more rapidly, from fleeting messages or briefly presented information.

Before considering these claims, some features of the data should be noted. None of the studies used very large or representative samples of subjects — the largest, a study by Nettlebeck and Lally[2], used 48 subjects and found a correlation of -.80. The Edinburgh studies have all been quite small and several have included mental retardates who may have contributed disproportionately to the correlations since it is difficult to find any psychological task on which they do not perform poorly. There is some disagreement about whether inspection time is more strongly related to verbal or performance IQ. One would like to see a study of a large, representative sample of the population with normal mean and distribution of IQ, which employed factor analysis on a number of different IQ tests in order to discover which component or components were most closely related to inspection time. These reservations are not trivial, but they are not significant enough to discredit what remains an intriguing and possibly very important finding. What does it mean?

One rationale for the search for simpler correlates of intelligence is the belief that they may reflect the basic psychological mechanisms underlying the diversity of intelligent behaviour. Such a belief antedates the development of IQ tests themselves. In the 1870s, some 25 years before Alfred Binet invented the intelligence test, the early exponents of the science of experimental psychology founded by Wilhelm Wundt at Leipzig studied such supposedly elementary processes as sensory thresholds, reaction times and digit spans. Some experimenters tried to demonstrate that differences in mental ability could be reduced to differences in these elementary processes.

They were not very successful. A celebrated study of students at Columbia University showed that these sorts of measures correlated neither with one another nor with such an obvious (albeit imperfect) criterion of intelligence as a student's grades. What was to be made of this? According to Binet, the answer was to abandon psychological theory and to take a more pragmatic approach. If we want to measure a child's intelligence then we should ask him to solve simple problems, answer questions about the meanings of words, or show that he has understood a passage of prose read out to him. No matter that we have no theory of the nature of the psychological processes involved in such tasks. What counted was that the tests worked: not only did they correlate with one another, but they also served to differentiate older children from younger ones and those judged more clever from those independently judged less able. The intelligence test, later to be called the IQ test, was born and flourished on this purely pragmatic basis. Items were selected for inclusion in IQ tests not on the basis of any theory of the mind, but because they discriminated between children of different ages or more simply because they agreed with other items, and because the new test agreed with the old.

This extraordinary absence of underlying theory has always been the Achilles' heel of IQ tests. No matter how well their tests predicted success in school, university or the world at large, no matter that factor analysis revealed that all tests were measuring a single underlying trait of general intelligence or '*g*', it was always possible to embarrass psychometricians by asking them just what was this general intelligence that they were measuring. The fact is that they had no idea, and no talk of 'innate general cognitive ability' by Burt, or of the 'eduction of correlates' by Spearman could disguise this fact. It is small wonder that many differential psychologists should have sought some more substantial psychological foundations on which to build their tests. Both Eysenck[3] and Jensen[4] revived the idea that differences in speed of decision-making, as evidenced in choice reaction times, may underly differences in intelligence. When a subject is presented, on each trial, with one of a set of possible stimuli and must respond appropriately, but as quickly as possible, it has long been known that his reaction time usually increases with the number of alternative stimuli. Correlations of between -.30 and -.50 have been reported between IQ and the slope of reaction time against number of alternatives.

But these correlations, although often significant, are not nearly as impressive as those reported by Brand and by Nettlebeck and Lally. It is possible, as Brand would argue, that reaction times measure other processes which serve only to obscure the fundamental relationship between intelligence and speed of perceptual processing.

Can intelligence really be reduced to such a basic process as this? Surely intelligence is too diverse to be captured thus simply. To this objection, differential psychologists have a ready answer. For all its apparent diversity, they argue, intelligence is a unitary trait, and IQ tests have revealed it to be so. As Clarke and Clarke[5] point out, "the most outstanding feature of all measures of intellectual ability in its broadest sense is that they show some degree of positive covariation". The simplest explanation of this covariation is that all IQ tests measure, with varying degrees of accuracy, this single trait of general intelligence or *g*. And the newly discovered correlation between IQ and inspection time reveals that *g* is a matter of speed of taking in perceptual information.

The next, marginally more sophisticated objection must be that the theory rests a great deal of weight on a set of correlations. The facts reported are consistent with the theory, but correlational data always admit a variety of explanations. That most tests of mental ability show marked positive correlations with one another is

---

N.J. Mackintosh is a Professor in the Laboratory of Experimental Psychology and Centre for Research on Perception and Cognition at the University of Sussex.

0028-0836/81/070529-02$01.00

important and perhaps surprising (slightly less so when one remembers that one criterion for the inclusion of items in IQ tests is that they should correlate with other items already in the test). But it cannot prove that all items are measuring the same underlying trait. There might be a variety of independent traits, which happened to correlate with one another in the general population. There might be an indefinitely large number of traits, skills, strategies, or bits of knowledge possessed by the intelligent man, and different items in IQ tests might sample different subsets of the available pool. Positive correlations between different items or tests would reflect only the overlap between the subsets sampled by each. The postulation of *g* is only one of several possible explanations of the observed pattern of inter-correlations, which is why it may be important to provide independent validation for the construct. But is it not obvious that the discovery of some new characteristic, however simple, basic, or well defined, which correlated with IQ would achieve a great deal? A correlation between IQ and inspection time does not mean that IQ *is* mental speed, nor even that speed is one (of several) causes of IQ. The causal relationship might be very much less direct than this; it might even go in the reverse direction.

If we are to understand the relationship between IQ and inspection time, we shall need more help from theoretical psychology, in particular from workers in cognitive psychology and artificial intelligence. Unfortunately, psychometricians have looked to neurophysiology for the support they perceive is needed, and even more unfortunately to bad neurophysiology. Hendrickson and Hendrickson[6] for example, have claimed to show that IQ correlates with measures of average evoked potentials, and have proposed an implausible theory of cerebral function to explain the correlation. The implication of their argument is clear: here is a correlation where there can be no question about the direction of causality: differences in intelligence must depend on differences in the efficiency of neural activity. In a very general way, they may be right — somehow, no doubt, intelligence depends on the brain. But in our present ignorance of the neural mechanisms underlying any aspect of cognition, specific hypotheses are likely to seem naive, and the Hendricksons neurophysiology and neurochemistry will seem to most specialists something worse than this.

Differential psychologists have seen a second advantage to be gained from the discovery of a correlate of IQ such as inspection time, reaction time or evoked potentials. They are sensitive to the charge that IQ tests are biassed. Differences in IQ between middle and working classes and between whites and blacks in the United States are well established, but are widely seen as evidence only of the bias inherent in the tests. The charge may not be as strong as some critics would like to believe, but it will always seem plausible that tests of vocabulary, information, arithmetic, or even simple problem solving may discriminate unfairly against those whose background has made them less familiar with the material in question. But who could claim that a test which asked one to say which of two briefly exposed lines was the longer could be culturally biassed or loaded against minority groups? Here, at last, is the psychometrician's dream: a culture-free test of intelligence.

The argument may seem persuasive, but a moment's reflection should give one pause. For suppose that the results of such a test revealed as large a difference between North American blacks and whites as is revealed by current IQ tests. There is no reason to believe that this would persuade the critics of IQ tests that blacks really were, on average, less intelligent than whites. It would be just as likely to undermine the credibility of inspection time as a measure of intelligence. If, on the other hand, such a test revealed no average difference between blacks and whites, this would not prove that conventional IQ tests were biassed and that there really were no differences in intelligence, for that would only be true if we were willing to accept inspection time as the better measure of intelligence. And since, *ex hypothesi*, its correlation with IQ would now have been degraded, one might wonder how someone who puts great store by the importance of these correlations in the first place could reasonably advance such a proposition.

The discovery that IQ correlates very highly with performance on an apparently simple psychological task, although intriguing, probably important, and surely destined to instigate some valuable research, does not of itself resolve some of the psychometricians' problems. It does not justify Eysenck's assertion[7] that "intelligence can now be measured by the methods of natural science — it is not just a mythical or abstract thing". As Eysenck has often argued, measurement is theory based. The theory relating IQ and inspection time seems distinctly simple. It may be correct, but most cognitive psychologists would be inclined to agree with Seymour and Moir[8] that "qualitative variations in the executive programmes which schedule and call the subordinate routines are more important as cognitive determinants of intelligence than differences in the speeds of functioning of the individual routines." □

---

1. Brand in *Intelligence and Learning* (eds Friedman, Das & O'Connor) (Plenum, New York, 1980).
2. Nettlebeck & Lally *Br. J. Psychol.* **67**, 17 (1976); *Am. J. ment. Defic.* **82**, 273 (1977).
3. Eysenck *Br. J. educ. Psych.* **37**, 81 (1967).
4. Jensen *Bias in Mental Testing* (Methuen, London, 1980).
5. Clarke & Clarke in *Developmental Psychology* (eds Sants & Butcher) (Penguin, London, 1972).
6. Hendrickson & Hendrickson *Personality and Individual Differences* **1**, 3 (1980).
7. Eysenck *Intelligence and Learning* (eds Friedman, Das & Connor) Plenum, New York, 1980).
8. Seymour & Moir *Br. J. Psychol.* **71**, 53 (1980).

The source of enduring meteor train luminosity

from W. J. Baggaley

In the second issue of *Nature* (**1**, 58; 1869) the journal's founder Editor, Norman Lockyer, commenting on a first report of an hour-long luminous meteor train, hoped "... that advantage was taken of this almost unprecedented opportunity to bring the spectroscope to bear upon a meteor cloud ..." and expressed the wish that the mechanism of this enigmatic phenomenon be unravelled. More than a century later spectral observations of long-enduring meteor trains are still few and our understanding of the processes that produce the persistent luminosity is still incomplete. The recent measurement (Hapgood *Nature* **286**, 582; 1980) of emission in the near infra-red of a 32-minute train is therefore a substantial contribution to the meagre data.

The problem is that enduring trains are rare; the extensive survey of observations of both sunlit trains occurring at twilight and night time self-luminous trains compiled by C. P. Olivier (University of Pennsylvania) shows that trains longer than ten minutes occur in about one per 125,000 visual meteor events. The corresponding occurrence frequency for a single observer is therefore about 8×10^{-5} h^{-1}. Since a train may occur in any part of the sky and may have an unfavourable orientation, the problem of securing spectral records is clearly formidable. Spectroscopic observations of both the meteor primary (impact excitation) and the enduring emissions were undertaken in the last century using multi-prism direct-vision slitless instruments. The first observations of enduring trains, by Herschel in 1866, were of the Perseid shower. He and later workers reported the regular appearance in trains lasting minutes of yellow and green emissions identified as the Na D doublet at 5,893 Å and the Mg triplet at 5,183 Å as well as bands in the orange and red. The only other visual spectrum obtained appears to be the transmission grating observation reported by Fred Whipple (in *Rocket Exploration of the Upper Atmosphere* eds Boyd & Seaton, 1954) of an orange-red band

---

W. J. Baggaley is in the Department of Physics, University of Canterbury, Christchurch, New Zealand.

0028-0836/81/070530-02$01.00

accompanied by a discrete red emission. Multiple-filter visual binocular studies (Hajdukova *Bull. astr. Insts Csl.* **18,** 187; 1967) demonstrated that for a non-shower sample of 114 trains radiation was predominantly longward of about 5,500 Å, a result in contrast to primary emission which is due mainly to features in the blue. Since primary spectra of meteors are dominated by short-wavelength iron transitions, the early orthochromatic emulsions were successful in white light photography. However, attempts to record enduring trains were largely unsuccessful until the use of panchromatic emulsions extended the spectral response to about 6,500 Å. More recent observations (Spalding & Hemenway *Astrophys. J.* **66,** 54; 1961) suggested the presence of a narrow spectral region in agreement with the photographic results (Bakharev *et al. Bull. Inst. Astrophys. Acad. Tadzik, SSR* **53,** 14; 1970) of 15-second trains which showed an absence of blue radiation with contributions predominantly between 5,000 and 6,000 Å.

Photograph of a meteor trail in close proximity to the Andromeda Galaxy M31. (Courtesy of Ondrejov Astrophysical Observatory, Czechoslovak Academy of Science.)

In the latest observation, Hapgood measured in the passband between 7,000 and 9,000 Å an emission rate per unit length of the meteor column of 2×10^{17} photon s^{-1} m^{-1} and attributed the emission to the excitation of 'forbidden' oxygen states O_2^* during the reduction of ground-state sodium monoxide in the cyclic mechanism

$$\mathrm{Na}+\mathrm{O}_3 \rightarrow \mathrm{NaO}+\mathrm{O}_2$$
$$\mathrm{NaO}+\mathrm{O} \rightarrow \mathrm{Na}+\mathrm{O}_2^*$$

The present picture therefore is that enduring trains radiate predominantly at wavelengths greater than about 5,000 Å and the evidence suggests that the process is mainly band emission. Indeed, for the abundant meteor species (Fe, Mg, Ca, O, Na) there are no strong low-energy allowed atomic transitions in the red and IR. In any reactant-consuming emission process resulting from an interaction between the species of a meteor column or between column species and atmospheric constituents, the emission rate will vary approximately inversely as the meteor column cross-sectional area. Hence under the action of diffusive (ambipolar or molecular) expansion such luminosity could not survive for more than a fraction of a minute. The production of train luminosity for tens of minutes requires a high photon yield process with regeneration of an active species, such as the sodium catalytic reaction cycle. This oxidation–reduction process invoked to explain enduring meteor trains has recently gained unequivocal support as a source of the sodium nightglow.

Hapgood proposes that excitation of the atmospheric system during the reducing step produces the IR emission. The formation of vibrational levels of the second excited electronic state of molecular oxygen $O_2(b^1\Sigma_g^+, v)$ by ground-state reactants NaO and O is not in violation of energy or spin conservation or of correlation rules and an appreciable fraction of reducing reactions might be expected (Bates & Ojha *Nature* **286,** 790; 1980) to yield O_2 in either the $b^1\Sigma_g^+$ or $a^1\Delta_g$ state. Unfortunately, any attempt to gain information by airglow studies on the channelling of reaction energy into the O_2 $(b^1\Sigma_g^+)$ state is not possible because oxygen atom recombination processes swamp any contribution from the Na catalytic mechanism.

While there is support for the oxidation–reduction cycle as the origin of the sodium nightglow and as a contributor to enduring train luminosity (the only other meteoric species for which the cyclic process is exothermic in the production of excited atoms is potassium with the possible yield of the 7,665 Å doublet), the role of the Atmospheric system must remain conjectural. Certainly the Atmospheric bands cannot account for the visual characteristics of trains — the green, red and orange emissions. This deficiency can be filled, however, by cyclic processes known to be exothermic involving metal and alkaline earth metals where regeneration of active species Fe, Al, Ca and Mg — atoms deposited during meteoroid ablation — may lead to high-yield monoxide band emissions and IR continua. In recent years laboratory studies of such chemiluminescent reactions have employed low-temperature flames and flow reactors using controlled atmospheres with various oxidizing and reducing agents as well as cross-beam techniques where processess under single collision conditions can be monitored. Such laboratory work has been complemented by photometric studies of chemiexcitation emissions resulting from rocket releases of various reagents into the thermosphere. Thus, excitation of diatomic FeO in chemiluminescent reactions with ozone oxidation producing the orange systems with strong emissions 5,500–6,500 Å has been studied in the laboratory while the bands have been observed in night-time atmospheric ejections of iron carbonyl, $Fe(CO)_5$. The blue-green system of AlO, has received much attention in the laboratory and numerous atmospheric releases of trimethyl-aluminium have been monitored: the reaction cycle excites strongly many sequences 4,500-5,000 Å as well as a broad IR system. Similar laboratory work has examined both the excitation in CaO of the red system yielding strong emission in the range 5,981–6,289 A together with the IR system at 7,712–9,229 Å; and also the excitation in MgO of the green system where the transition sequence yields strong emission from 4,900 to 5,020 Å. These chemiluminescent reactions exhibit photon yields of about 10 per cent and also produce IR continua due to the excitation of polyatomic intermediaries.

The spectral features of the oxide–reduction cycle emission involving meteoric atoms correlate well with luminous train characteristics: the absence of short-wavelength radiation, the narrow band emissions in the yellow, orange, red and green (where the 100 Å-wide MgO band could be confused with the Mg 5,183 Å triplet in a low-resolution spectroscope), and the IR bands and continua. Metal oxide excitation is therefore a strong candidate for the source of enduring train luminosity.

Image intensifier photon integration systems with their sensitivity to faint trains can alleviate the problem of the paucity of meteor events and play a valuable role in the provision of definitive spectral data. Such data should provide unequivocal evidence for the source of enduring train luminosity and fulfil the wishes of the first editor of *Nature.* □

Selection and induction in the development of autonomic neurones

from Miranda Robertson

ONE of the most important recent advances in the study of neural development has been the demonstration by biochemists, physiologists and electron microscopists at the Harvard Medical School that dissociated neurones from the sympathetic ganglia of newborn rats can be induced to switch their neurochemical character by manipulation of the conditions in which they are cultured[1]. Cultured neonatal cells at first synthesize and secrete noradrenaline, but can be induced to switch to the synthesis and secretion of acetylcholine by co-culture with non-neuronal cells such as heart, or by medium in which such non-neuronal cells have been grown. At the same time, N.M. Le Douarin and her colleagues at the CNRS laboratory in Nogent-sur-Marne have demonstrated neurochemical lability in avian sympathetic ganglion cells *in ovo*. They have been able to induce adrenergic function in cholinergic sympathetic ganglia of quails by back-transplanting them into the neural crest of earlier embryos at a site from which adrenergic ganglia normally develop[2]. These discoveries suggested that neurochemical phenotype might be determined after neuronal differentiation by tissues in the environment, and stabilized in the course of maturation[3].

In the light of more recent developments, however, it seems possible that neurochemical phenotype, at least in the autonomic nervous system, may never be irreversibly stabilized. This possibility was extensively discussed at a recent conference*, with two other closely related issues: what is the primordial phenotype of the autonomic neurone — adrenergic, cholinergic or dual-function; and are the switches reported *in vitro* and *in vivo* all due to induction, or may some of them actually reflect selection from an already heterogeneous population?

There is now evidence that although cultures of neonatal rat ganglion cells, at least in the hands of the Harvard team, express adrenergic characteristics, chick neural crest cells are already able to express the cholinergic enzyme choline acetyltransferase (CAT) at the point at which they begin to migrate to form the ganglia (Le Douarin). Unless there is a fundamental difference between chick and rat, therefore, the Harvard dissociated neonatal neurones may already have a cholinergic history.

Furthermore, some recent experiments by M. I. Johnson (Washington University) imply that they may never entirely abandon their cholinergic option. Dissociated neurones (though not, for unknown reasons, explants) even from adult rat superior cervical ganglion can be induced to express cholinergic functions in appropriate conditions *in vitro*. This important result lends weight to the view, held by several participants at the meeting, that autonomic neurones may remain inherently dual-function throughout life, their detectable phenotype depending on the balance of environmental factors.

The alternative view, that the apparent switch in neurotransmitter specificity may be due not to induction but to selection of already committed neural cells from a heterogeneous population, was persuasively put by D. Edgar, from Thoenen's laboratory at the Max-Planck-Institut, Martinsried. He and his colleagues have identified three factors — nerve growth factor (NGF) and factors from medium conditioned by heart cells or glial cells — which are required by different subpopulations of chick autonomic ganglion cells at different developmental ages for survival *in vitro*. They have now been able to show that the same subpopulations can also be distinguished by markers of adrenergic or cholinergic functions[4].

Does this mean that the actual effect of a change in culture conditions, at Harvard and St Louis as well as at Martinsried, is to enhance the growth of a separate subgroup of existing cholinergic cells at the expense of the adrenergic population, rather than to induce a switch? The answer to this question must be no. The Harvard team, represented at the meeting by D. Potter, has been able to show that single autonomic neurones in culture can, both on the physiological criterion of their ability to excite or inhibit cardiac contraction and on the ultrastructural criterion of vesicle morphology, switch from adrenergic to cholinergic. The St Louis team (in whose hands ganglion cells do not switch completely but remain dual-function) have been able to confirm immunocytochemically the expression of adrenergic traits in cells shown electrophysiologically to display cholinergic properties.

How, then, can the different results be reconciled? Edgar suggests that the explanation may lie in the relative efficiency of the culture techniques used in the different laboratories. Perhaps only some cells are capable of switching, and those cells are selected by the dissociation and plating procedures used at Harvard. This view is consistent with the fact that plating efficiency at Harvard is about 10% whereas at Martinsried it is about 90%. The implication is that the Martinsried cultures are more representative of the original cell population; but does it follow that their behaviour is representative of what actually happens *in vivo*? So far, with one possible exception, there are no data which can be unequivocally interpreted one way or the other. The exception is the investigation recently reported by Landis at Harvard on the sweat glands of the rat footpad, which seem to be innervated early in ontogeny with adrenergic nerves that subsequently become cholinergic[5].

Nonetheless, there can be no doubt that survival factors such as NGF play an important part in shaping the developing nervous system; and I. A. Hendry (Australian National University) has confirmed that NGF injected into the rat eye is transported back into the superior cervical ganglion and can reverse the effects of axotomy, which would normally cause the neurone to die. This implies that neurones depend for survival on factors retrogradely transported from their target tissues, although definitive evidence would of course require that the NGF be endogenous and not injected.

A more characteristically ambiguous discovery was reported by I. B. Black (Cornell University), who finds cells staining with catecholamine immunofluorescence in various ectopic sites in rat embryos treated *in utero* with NGF. This result can be interpreted either as induction of catecholaminergic characteristics in cells which would not otherwise express them (or would express them at levels too low to be detected by Black's technique); or as rescue of cells dependent on NGF and which would normally be deprived of it and die.

What are the limits on the ability of a ganglion cell to switch according to its environment? Le Douarin finds that although embryonic sensory glia can give rise to glia, neurones and endocrine cells of the autonomic nervous system, transplanted autonomic ganglion cells seem to be committed to an autonomic fate, and invariably migrate to autonomic ganglia, showing that the range of their developmental potential is narrower than that of neural crest or even sensory ganglion cells.

Furthermore, Potter pointed out that his experiments with Furshpan and Landis on single cultured neurones followed over time seem to show variation from one neurone to another in the extent of the adrenergic-to-cholinergic switch, some cholinergic neurones retaining catecholamine uptake and storage mechanisms while others seem to lose them. While some of the variation seems to be intrinsic to the individual neurone, Potter emphasized that the expression of different neurotransmitter-specific characteristics is

Miranda Robertson is associate editor of Nature.

---

*Ciba Foundation Symposium on 'Development of the Autonomic Nervous System' held in London from the
1980.

0028-0836/81/070532-02$01.00

differentially sensitive to changes in the culture medium. Defining the properties of the medium, or the dosage of 'factors' responsible for the induction of the different traits is one of the most important aims of current research *in vitro*.

Participants at the meeting hoped that in the longer term, their research might elucidate diseases of the autonomic nervous system: for example, it seems possible that familial dysautonomia may be due to a deficiency in NGF, and that Hirschprung's disease may be due to the failure of ganglion cells to migrate into the hindgut. At this stage, it seems too much to hope that fundamental research will produce clear clinical answers; but there is no doubt that it is asking the relevant questions. ☐

1. Patterson, P. H. *A. Rev. Neurosci.* **1**, 1 (1978).
2. Le Douarin, N. M. *Nature* **286**, 663 (1980).
3. Walicke, P. A., Campenot, R. B. & Patterson, P. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5767 (1977).
4. Edgar, D., Barde, Y.-A. & Thoenen, H. *Nature* **289**, 294 (1981).
5. see Robertson, M. *Nature* **288**, 533 (1980).

Quasar alignments: chance or necessity?

from M. G. Edmunds

EVEN the staunchest believers in the interpretation of quasar redshifts as indicating their extreme distances from us, must be experiencing some disquiet at the apparently extraordinary configuration of quasars recently reported by H. Arp and C. Hazard (*Astrophys. J.* **240**, 726; 1980).

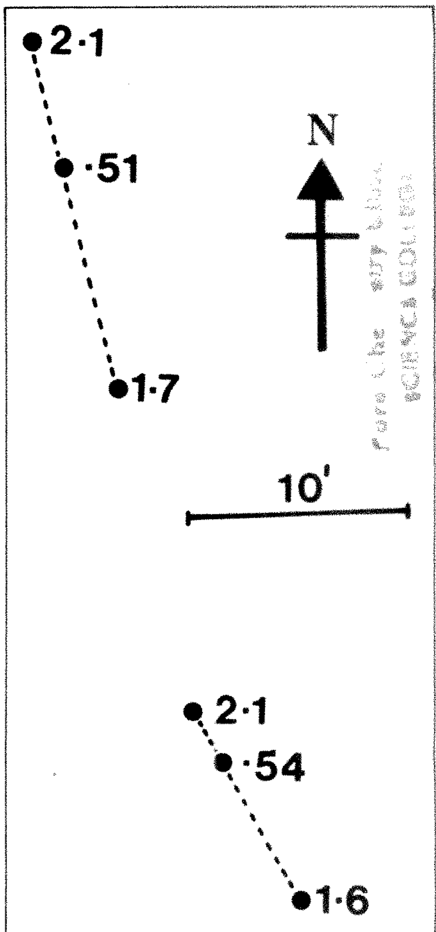
What Arp and Hazard have found is shown schematically in the accompanying figure. Two triplets of quasars are quite close together on the sky, with each triplet lying in a straight line which is exact to within the image sizes (about 2 seconds of arc) despite overall quasar separations of some 600 seconds for one triplet and nearly 1,000 seconds for the other. The confirmation of the identification of the objects as quasars, and the determination of their redshifts, have been accomplished by spectroscopic observations. The quasars are not unduly faint, all being brighter than about twentieth magnitude in visible light. Recent surveys of the density of quasars on the sky (see, for example, A. Savage and J. G. Bolton *Mon. Not. R. astr. Soc.* **188**, 599; 1979) imply that roughly five to ten quasars would be expected down to this magnitude per square degree of sky, so this peculiar configuration would seem, at first sight, to be highly unlikely to occur by a chance positioning of the individual objects in the sky.

The obvious implication is of association of the objects in the triplets, although the differences in redshift rule out any kind of gravitational lensing of a single object. But a kinematical explanation can be proposed, where the outer quasars in a triplet have been thrown out in opposite directions at relativistic speeds by the central object. This is not a new idea, but it has been unpopular as an explanation for quasars since it would be expected that some objects should be thrown out towards Earth giving Doppler blueshift, and blueshifted quasars are never observed. A novel mechanism for suppressing blueshifts has been suggested by Sir Fred Hoyle in a recent preprint (erroneously titled *The Quasar Controversy Resolved,* since it is bound to provoke even more discussion!). By assuming that the quasars ejected from the central object only radiate backwards in their rest frames into a cone around the line defined by the motion of the ejected object through a stationary intergalactic medium, then the so-called relativistic 'headlight' effect can cause this radiation to beam forwards towards an observer on Earth, while the transverse Doppler effect will cause it to be redshifted. Provided the cone of emission has suitable angle, blueshifts need never be observed. Unfortunately Sir Fred does not apply the model in detail to Arp and Hazard's triplets, and it is difficult to judge immediately just how artifical the parameters might have to be to give the observed configuration.

The idea of ejection of energetic objects from galactic (or quasar) nuclei is not in itself particularly heretical. The classic twin-lobe structure of giant radio sources, frequently observed to have a galaxy or quasar at the centre, certainly implies the symmetric ejection of something, although it could be just radiation. It is interesting that searches for optical objects at the regions of highest emission in the radio lobes, the very place you might expect ejected objects, has not found anything looking like a bright quasar. Some optically emitting regions have been found (see W. C. Saslaw, J. A. Tyson and P. Crane *Astrophys. J.* **222**, 435; 1978) but they are far too faint for reliable spectroscopic examination.

Is it possible that there are two kinds of quasar — one set composed of very active galactic nuclei, extremely energetic and at cosmological distances, and another kind composed of intergalactic ballistic missiles? This is a problem which is not new to the pages of *Nature* (M. Rowan-Robinson, **262**, 97; 1976; G. Burbidge **282**, 451; 1979), and it would be difficult not to concede that at least a proportion of quasars are at cosmological distances. Perhaps the most compelling arguments for this are the apparent continuity of spectral and energetic properties between Seyfert galaxies and quasars, and the observation of quasars which are apparently members of clusters of galaxies with the same redshift, implying a common (large) distance.

If some quasars are indeed distant, then there exists a theoretical problem in explaining the generation of their high luminosities in what must be (on the basis of observed variability) a comparatively small emitting region. Considerable progress has been made in recent years in considering accretion of material onto a massive black hole as a power source. Apart from the triplet alignment problem, the only real advantage in a ballistic quasar theory was that the objects could be in the local region near our own Galaxy and hence not require large energy outputs. Since the energy problem must be faced for



Schematic diagram of the configuration of quasars found by Arp and Hazard. The figures indicate the redshifts of the quasars, and it will be noticed that the two triplets have curiously similar values. The horizontal bar indicates the scale of ten minutes of arc on the sky.

M. G. Edmunds is in the Department of Applied Mathematics and Astronomy, University College, Cardiff.

the first type of quasar, it is of doubtful theoretical value to have a second type for which these problems can be sidestepped. Also, the overall similarity of spectral properties of quasars suggests a single kind of object. It is therefore tempting to apply Occam's razor and propose that there are only cosmological-redshift quasars, and that ballistic quasars do not exist. How could Arp and Hazard's alignments then be explained?

The answer could be that such associations are not as unlikely by chance as it might seem at first sight — perhaps the right statistics have not yet been done — it is anyway very difficult to apply *a posteriori* statistics once a peculiar pattern has been found. At least this point can, and should, be investigated more carefully by proponents of both quasar models.

It is perhaps surprising that the distances of quasars remain uncertain almost twenty years since their high redshifts were recognized, but a glance at the history of the debate over the nature of spiral nebulae should be enough to convince anyone that important questions sometimes take a long time to settle. Whatever their explanation, Arp and Hazard's fascinating triplets surely deserve a constellation name — Cosmology's Damocles or Hoyle's Banderillas perhaps? □

Antimatter back to front

from C.H. Llewellyn Smith

BEFORE 1956 it was taken for granted that parity is conserved — in other words, that the laws of nature do not distinguish between left and right. The shock produced by the discovery of parity violation was alleviated by Landau's observation that symmetry under the combined operation of 'charge conjugation' (C), which turns particles into antiparticles, and parity (P) could remain exact, that is, that there could be symmetry between left-handed matter and right-handed antimatter. If this CP invariance holds, then according to Landau "space remains completely symmetrical" and an asymmetry which "in view of the isotropy of space (conservation of momentum) would be more than strange" could be avoided. This view was generally accepted and CP invariance came to seem almost self evident. The discovery of CP violation by Christenson, Cronin, Fitch and Turley in 1964, for which they received the Nobel prize in 1980, was therefore almost as great a shock as the discovery of parity violation.

The existence of P and CP violation makes it possible to tell from watching a film of submicroscopic processes whether they involve matter or antimatter and whether the film is being projected correctly or has been loaded back to front. In contrast, if CP were conserved it would be impossible to distinguish a correctly shown film of matter from a back-to-front film of antimatter. If both C and P were conserved we could distinguish neither a film of matter from one of antimatter nor whether it had been loaded back to front.

Up to now CP violation has only been observed in the decays of neutral K mesons. The particles have fascinating properties which directly demonstrate fundamental aspects of quantum mechanics. In energetic collisions of strongly interacting particles, such as protons, either a K°, which contains an antistrange quark, or its antiparticle $\bar{K}^\circ$, which contains a strange quark, can be produced. However, strange quarks are unstable and decay into non-strange quarks. This allows K° and $\bar{K}^\circ$ to decay into identical products (either two or three pi mesons). It also allows a K° to turn into a $\bar{K}^\circ$ and vice versa. Consequently it is possible to make coherent quantum mechanical superimpositions or mixtures of K° and $\bar{K}^\circ$. Two particular mixtures behave as particles with definite masses and lifetimes. One, called K_S, is relatively short lived and decays predominantly into two pi mesons. The other, K_L, is longer lived and decays mainly into three pi mesons. A K° or $\bar{K}^\circ$ produced in a proton collision can be regarded as a mixture of K_S and K_L; the K_S component decays relatively quickly leaving pure K_L.

If CP symmetry were exact, the decay of K_L to two pi mesons would be completely forbidden, because K_S and the two-pion state into which Ks can decay would be symmetric under CP symmetry whereas K_L would be antisymmetric. The Cronin and Fitch experiment was designed to look for the CP violating decay $K_L \to \pi^+\pi^-$ and convincingly showed that it accounts for about two K_L decays in every thousand. A more direct manifestation of CP violation was provided by the later observation that the rates for the decays $K_L \to \pi^+ e^- \nu_e$ and $K_L \to \pi^- e^+ \bar{\nu}_e$ differ by about half a per cent; exact CP symmetry would turn K_L into itself, π^+ into π^-, e^- into e^+ and ν_e into $\bar{\nu}_e$ and require the rates to be equal.

Immediately after Christenson *et al.* announced their result in 1964, various ingenious ways to reconcile it with CP symmetry were proposed. Perhaps the most interesting was the idea of a long-range force which couples to K° and $\bar{K}^\circ$ and to other forms of matter and antimatter with opposite signs. Such a force would generate apparent CP violation due to the predominance of matter over antimatter in the Universe. However, it would also give rise to an apparent violation of lorentzian invariance causing the fraction of decays of K_L to $\pi^+\pi^-$ to depend on velocity. This was quickly ruled out.

It soon became apparent that not only CP but also time reversal invariance (T) is violated or, in other words, that the laws of nature depend on the arrow of time. This makes it possible to tell whether a film of certain submicroscopic processes is being run forwards or backwards. The possibility of T violation was immediately obvious because it was known that any theory which respects lorentzian invariance and the general principles of quantum mechanics is necessarily symmetric under the combined operation CPT. This CPT symmetry makes it impossible to distinguish a correctly projected film of matter run forwards from a back-to-front film of antimatter run backwards. If CPT is unbroken, then CP violation necessarily requires T violation. However T violation can be demonstrated directly without invoking CPT symmetry. Further study of the neutral K system showed that the quantum mechanical probability amplitude for K° to turn into $\bar{K}^\circ$ differs from the amplitude for $\bar{K}^\circ$ to turn into K° by about one per cent — a direct manifestation of T noninvariance. However, the amplitudes for K° and $\bar{K}^\circ$ to remain themselves, which must be the same if CPT symmetry holds, are equal to better than one part in 10^{17}.

It is hardly surprising that the origin of CP violation was mysterious in the 1960s since the nature of the weak interactions was not understood then. The advent of unified electroweak theories did not help at first since the simplest model which fitted the known facts did not allow CP violation. However, in 1973 Kobayashi and Maskawa pointed out that if there were more than four quarks CP violation would be almost unavoidable. Since the discovery of the fifth quark in 1977, CP violation has therefore seemed very natural. In fact, available unified models give rise to CP violation of the observed order of magnitude, although the precise strength of the effect can not be predicted at present. According to these models the only other likely observations of CP violation in the near future will be in the decays of mesons which contain the fifth quark and possibly also in neutrino oscillations, if they really occur, which are similar to K°- $\bar{K}^\circ$ oscillations.

This relatively happy situation has been jolted by the discovery that quantum chromodynamics (QCD), an extremely elegant and plausible candidate theory of the strong force, can give rise to CP violating effects, which are potentially much too large. Ironically one of the arguments for taking QCD seriously in the first place was that it could never generate strong CP violation. However, although this argument is true to all orders if the theory is treated perturbatively, it turns out that non-perturbative effects (called

---

C. H. Llewellyn Smith is a Reader in Theoretical Physics at the University of Oxford.

0028-0836/81/070534-02$01.00

'instantons') generate T and P violating effects which involve a new parameter called θ. Although θ could have any value, it would seem natural for it to be of order one. However, the experimental upper limit on the T and P violating electric dipole moment of the neutron requires $\theta < 10^{-8}$. No good reason why θ should be so small is known.

Thus there are still some mysteries surrounding CP violation. Indeed, it is still not clear whether C, P and T are intrinsically broken. It is possible to construct electroweak theories in which the underlying equations are completely symmetrical but the solutions are asymmetrical. Recently interest in the nature of CP violation has been greatly increased by the discovery that it may be a key element in explaining the large excess of matter over antimatter in the Universe. Attempts to unify the strong with the electroweak interactions suggest that baryon number is not conserved. Indeed, they predict that the proton decays with a lifetime of about 10^{30} years, and experiments designed to search for proton decay are now under construction. In this case, as originally discussed by Sakharov, a universe containing mostly matter can develop from a state of equilibrium between matter and antimatter because of the asymmetry due to CP violation. Furthermore, this picture is capable of explaining the observed ratio of about 10^8 between the number of photons and baryons in the Universe. If baryon number is conserved, it is necessary to invoke an initial asymmetry of one part in 10^8 to explain this ratio. If these cosmological speculations hold up they will put severe constraints on the fundamental mechanism responsible for CP violation.

The discovery of CP violation completed the conceptual revolution begun by the discovery of parity violation. This revolution has led to a healthy scepticism about other old shibboleths such as baryon conservation. Current theory can naturally accommodate CP violation. It is true that its magnitude cannot be predicted at present but this problem is on the same footing as other unanswered questions, such as how to predict the masses of quarks, if we leave aside the θ problem. The asymmetry between matter and antimatter which previously seemed "more than strange" now seems natural and perhaps even inevitable. But without it the Universe might be quite different. □

Genetic engineering and foot and mouth disease vaccines

from J. B. Brooksby

WHEN genetic engineering was first discussed in relation to foot and mouth disease vaccines, conventional virologists were sceptical of a practical outcome in less than perhaps ten or fifteen years. Enough was known of the viral genome to suggest that appropriate information might be incorporated into plasmids but a pessimistic view was taken of the possibility of expression of this information to provide the antigenic moiety of the virus in active form. The paper by Kupper and his co-authors on p.555 of this issue shows that the obstacle of expression has been surmounted and a milestone has been passed on the road to production of antigens for vaccine.

Although this is not the first report of the production of a viral antigen in this way, it is a happy circumstance that a foot and mouth disease product is early in the field, not perhaps because the agent was the first animal pathogen discovered to be a virus, but because of the importance of foot and mouth disease vaccine in controlling what is still one of the most serious diseases of animals throughout the world. More foot and mouth disease vaccine is produced than any other vaccine, but more is yet needed to begin campaigns in countries where the disease continues to spread. Compulsory regular vaccination in Western European countries over the past 25 years covering more than 60 million cattle has reduced the incidence of the disease to vanishing point — indeed, most recent outbreaks have been linked, at least circumstantially, to vaccine containing residual live virus. In South America, which has a long-standing problem, at least 500 million doses of polyvalent vaccine are produced annually. These figures indicate the importance of the disease to the animal industry, through its debilitating effects, lowering of milk production and interference, because of its remarkable spreading power, with trade in animals, meat and other products.

Present vaccines are produced from virus cultivated either in surviving tongue epithelial tissue collected from animals slaughtered for meat, or in cell lines grown in suspension in large (2,000 litres or more) vessels. Problems arise in the production of sufficient viral antigen, especially with some strains of virus, and the multiplicity of types (there are seven) and sub-types (over 60) necessitates either expensive polyvalent vaccines or vaccines appropriate to particular areas which may have to be changed rapidly to meet new strains arising in the field. Thus, although much has been achieved with available vaccines, as the experience of European countries shows, there is still a real need for alternative methods for producing viral antigen.

The application of genetic engineering may offer such a method and the work reported, together with other studies proceeding in the UK (at the Animal Virus Research Institute and Wellcome Foundation Ltd) and in the US (at the Plum Island Animal Disease Center and Genentech — see *Nature* **288**, 663; 1979), indeed give promise for the future.

The advantages of a vaccine based on an antigen consisting of the appropriate viral protein, prepared in a way that does not involve infective material at any subsequent stage after isolation of the viral RNA and which allows the production of greater quantities than from natural sources, are obvious. Kupper *et al.* have concentrated on the preparation of VP1, the coat protein responsible for the production of the antibody response to foot and mouth disease virus. Laporte and Bachrach and their colleagues in France and the US showed that purified VP1 from disrupted virus stimulates an antibody response, albeit weak, in guinea pigs and swine, and large doses will protect swine against subsequent infection.

In the present experiments, VP1 produced in the bacterial cell has been identified by its reaction with specific VP1 antibody and from the serological reaction it is estimated that at least 1,000 VP1 molecules are synthesized per cell. This is a signal achievement which is not diminished by the question — what remains to be done? Clearly, the new VP1 must be shown to immunize animals. The quantitative aspect is important, for the earlier work suggested that at least 1,000 times more antigen was required when the VP1 had been isolated than when it was carried on the intact, but inactivated virus particle. Yields from bacterial culture can probably be raised above those from mammalian cell culture and the factor by which this is possible will be critical for the success of the method. The quality of the antigen is also fundamental. It may be that, in order to preserve VP1 from degradation, it will be advantageous to synthesize the precursor which cleaves to form all four capsid proteins.

A new and exciting prospect is the possibility of overcoming some of the difficulties created by virus variation in the field. It may well be practicable to tailor strains to stimulate immunity to a wider antigenic spectrum than that given by naturally occurring viruses. There is no doubt that the progress reported will lead to a redoubled effort by the various groups and the prospects for new foot and mouth disease vaccines are indeed bright.

J. B. Brooksby is a former Director of the Animal Virus Research Centre, Pirbright, UK.

0028-0836/81/070535-01$01.00

Well-groomed predecessors

from R. D. Martin

RECONSTRUCTIONS of mammalian phylogeny have depended heavily on dental evidence, since teeth are preferentially preserved in the fossil record. As a rule anterior teeth (especially incisors) are easily lost in fossilization and most weight has been placed on the characteristics of cheek teeth (premolars and molars). However, there are some striking modifications of the lower anterior teeth among living mammals and these provide not only useful diagnostic features but also valuable functional clues. A good example is the 'tooth-comb' formed by the procumbent lower front teeth in mammal groups such as the tree-shrews, strepsirhine primates (lemurs and lorises) and 'flying lemurs' (which actually are not lemurs and do not fly — they are unusual gliding mammals relegated to their own order, Dermoptera). The tooth-comb of lemurs and lorises is particularly unusual in that the lower canines are incorporated along with the incisors, and it is now widely accepted that a unique six-tooth comb (two canines and four incisors) was probably an ancestral feature of lemurs and lorises.[1,2] In tree-shrews, by contrast, the tooth-comb is formed by some or all of the six lower incisors without involvement of the canine teeth, and in flying lemurs the four lower incisors all have crenulated tips and act as individual combs. These are doubtless convergent developments.

The tooth-comb has often been seen as an adaptation for grooming of the fur. Indeed, virtually all tree-shrew, lemur and loris species have been seen using the tooth-comb in a characteristic rake-and-lift action to groom the fur, though no such use of the lower incisors has been recorded for the flying lemur. Despite some past claims to the contrary[3,4] there is now no doubt that the tooth-comb serves a valuable grooming function in tree-shrews and in strepsirhine primates. Nevertheless, in recent years evidence has accumulated that all of these species also use their anterior lower teeth for feeding. For instance, soft fruit pulp may be scooped out with the tooth-comb and detailed field observations of nocturnal lemurs and lorises [5] have revealed that many species feed upon gums (natural polysaccharide exudates from trees). For one species — the lesser bushbaby (*Galago senegalensis*) — clear evidence of the use of the tooth-comb in gum-feeding has been collected[6]. Further, the greatest development of the tooth-comb (once body size has been taken into account) is found in two species which feed predominantly on gums, the needle-clawed bushbaby (*Galago elegantulus*) and the fork-crowned lemur (*Phaner furcifer*). It has been claimed[7] that the tooth-comb in small-bodied lemurs and lorises is too fragile to allow for its use in feeding, but fruit pulp provides no great resistance and gums are usually collected in a semi-liquid state. It is therefore a moot point whether grooming or feeding was the primary function of the tooth comb in lemurs and lorises, but it is certain that both functions are served in extant species.

Our understanding of the origin of tooth-comb grooming has now been considerably advanced by the report by Rose, Walker and Jacobs (see this issue of *Nature* p583) that wear patterns produced by repeated passage of hairs between the lower anterior teeth can be clearly recognised with the scanning electron microscope (SEM). SEM photographs of the lower anterior dentitions of extant strepsirhine primates, such as *Galago crassicaudatus*, reveal fine vertical grooves on the sides of the teeth. Similar grooves are found on the lower incisors of certain tree-shrew species, but no such wear patterns are found on the comb-like lower incisors of the flying lemur, suggesting that their function is restricted to feeding. This new evidence is particularly valuable because the findings can be extended back through the fossil record. A tooth-comb consisting of six lower incisors has been reported[8] for early Tertiary (Palaeocene/Eocene) arctocyonid condylarths and Rose *et al.* have demonstrated that the lower incisors of these early placental mammals bear grooves closely resembling those found on the comb-teeth of modern lemurs, lorises and tree-shrews. This neatly confirms the proposal made by Gingerich and Rose that the lower incisors of these condylarths were used for grooming. This would seem to be the earliest direct fossil evidence of a mammalian behaviour pattern, and it would certainly appear to be the most reliable demonstration that mammalian hair was definitely in existence over 55 million years ago! Significantly, the earliest (mid-Palaeocene) condylarth species examined by Rose *et al.* also exhibited wear on the tips of the comb-teeth, indicating a combined grooming and feeding function.

Reconstruction of the evolutionary history of the tooth-comb in lemurs is hampered by the poor fossil record. Until recently, the only reliable early fossil forms were early Miocene lorisids from East Africa (approx. 18-20 million years old). On the basis of indirect evidence, all three recognised East African lorisid genera from the Miocene (*Komba*; *Progalago*; *Mioeuoticus*) are thought to have possessed tooth-combs like their modern relatives, but no anterior tooth crowns are known[9]. Now Jacobs (see this issue of *Nature* p585) has reported new fossil finds from the late Miocene Siwalik deposits of Pakistan (some 7-10 million years old) assigned to the lorisid species *Nycticeboides simpsoni*. This newly-discovered species definitely had a tooth-comb, and Rose *et al.* have applied their SEM procedure to the crowns of one canine and two incisors. These teeth closely resemble their counterparts in modern lorisids and SEM examination revealed the fine vertical grooves characteristic of use in grooming. Thus, the use of the lorisid tooth-comb in grooming can be definitely traced back at least 7 million years. Unfortunately, though, feeding on gum or soft fruit pulp is not known to leave characteristic wear patterns on the tooth-comb, so possible dietary function remains untested.

Jacobs regards *Nycticebodies simpsoni* as belonging to the subfamily Lorisinae (including the slow-moving pottos and their relatives), rather than to the Galaginae (the agile, saltatory bushbabies). It is true that *Nycticeboides* shares with modern lorisines specific characters such as simple rear premolars, relatively weak development of the fourth cusp (hypocone) on the upper molars, and poor development of the ectepicondylar flange on the humerus. However, such characters only indicate a phylogenetic link between *Nycticeboides* and lorisines if they emerged *after* the divergence between galagines and lorisines (that is, if they are not primitive features of the lorisids), and this remains to be demonstrated. But if Jacobs is right, the presence of *Nycticeboides* in the Siwalik deposits could have a wider significance, since *Ramapithecus* (regarded by many palaeontologists as an early relative of man) also occurs in the Siwaliks. All modern lorisine species, unlike some bushbabies, are confined to relatively dense forests with virtually continuous arboreal pathways. Modern tree-shrews are similarly confined to forested regions and fossil tree-shrews have already been reported from these same Siwalik deposits[10,11]. The combined evidence of forest conditions provided by fossil lorisines and tree-shrews in the Siwaliks raises the possibility that *Ramapithecus* was also at least partly a forest-living primate. If *Ramapithecus* does lie close to the origin of the hominid line, as many authorities believe, confirmation of a forest background for this Miocene genus would be of great value, if only in ruling out some of the speculative suggestions which have been made in the past. □

R.D. Martin is Reader in Physical Anthropology, University College London.

1. Martin *Phil. Trans. R. Soc.* **264**, 295 (1972).
2. Cartini in *Phylogeny of the Primates* (ed. Luckett & Szalay) 313 (Plenum, New York, 1975).
3. Stein *Am. Nat.* **70**, 19 (1936).
4. Avis *Am. J. Phys. Anthrop.* **19**, 55 (1961).
5. Martin in *The Study of Prosimian Behavior* (ed. Doyle & Martin) 45 (Academic, New York, 1979).
6. Bearder & Martin *Int. J. Primatol.* **1**, 103 (1980).
7. Szaley & Seligsohn *Folia Primatol.* **27**, 75 (1977).
8. Gingerich & Rose *J. Mammal.* **60**, 16 (1979).
9. Walker in *Prosimian Biology* (eds. Martin, Doyle & Walker) 435 (Duckworth, London, 1974).
10. Chopra, Kaul & Vasishat *Nature* **281**, 213 (1979).
11. Chopra & Vasishat *Nature* **281**, 214 (1979).

0028-0836/81/070536-01$01.00

REVIEW ARTICLE

Dendritic release of dopamine in the substantia nigra

A. Cheramy, V. Leviel & J. Glowinski

Groupe NB—INSERM U 114, Collège de France, 11, place Marcelin Berthelot, 75231 Paris Cedex 05, France

Dopamine can be released in the substantia nigra from the dendrites of nigrostriatal dopaminergic neurones, to be involved there in the self-regulation of the dopaminergic cells, to control the release of neurotransmitters from nigral afferent fibres and to influence the activity of nigral non-dopaminergic cells.

IT is still widely considered that the transfer of information is unidirectional in neurones which communicate through chemical synapses. Messages travelling from dendrites, through somata and down axons finally reach the nerve terminals from which the neurotransmitter is released. By interacting with postsynaptic receptors located on somata or dendrites, the neurotransmitter will then modify the activity of the target neurones. Several years ago, however, it was suggested that molecules could also be released from dendrites to influence surrounding cells or nerve terminals[1-5]. In this article, we review the accumulating evidence which indicates that dopamine (DA) is released from dendrites of the nigro-striatal dopaminergic neurones and contributes to the processing of signals in the substantia nigra (SN). Complementary anatomical, electrophysiological and biochemical studies show that DA is released from dendrites to regulate not only the release of neurotransmitters from some nigral afferent fibres, but also the activity of the dopaminergic neurones and of other nigral efferent neuronal pathways. Consequently, by acting in the striatum and the SN, the dopaminergic neurones may influence the messages delivered to the striatum by the cortico- and thalamo-striatal fibres and those emerging from the SN through projections innervating the thalamus and the pontine nuclei.

Substantia nigra as a site for study of dendritic release of DA

Nigro-striatal dopaminergic neurones are particularly well suited to the study of neurotransmitter release from dendrites in the central nervous system. Their cell bodies are mainly concentrated in the pars compacta and their long, ramified and varicose dendrites extend throughout the pars reticulata of the SN[1]. Their fibres, in contrast to those of non-dopaminergic cells of the pars reticulata, do not emit collaterals within the SN[6]. Intraventricular or locally injected radioactive catecholamines readily label the cell bodies and the dentrites of the dopaminergic neurones whilst showing the virtual absence of catecholaminergic nerve terminals[7-10]. Furthermore, few if any nigral boutons degenerate after intraventricular or nigral injections of 6-hydroxydopamine[11,12], a neurotoxin which selectively destroys catecholaminergic fibres. The SN is also an appropriate site for the study of the dendritic release of DA because there is good evidence not only that DA can be synthesized[13-18] and stored[1,10,19,20,21,22] in dopaminergic dendrites, but that it can also be inactivated. However, it should be noted that the mechanism of DA re-uptake in dendrites may differ from that in nerve terminals[1]. Furthermore, little is yet known about the mechanisms of DA storage in dendrites. Although a few large or small dense core vesicles are seen to be scattered in some thin distal dendrites using 5-hydroxydopamine as a marker of monoamine storage sites[23], no specialized vesicles are seen in dopaminergic dendrites identified by the autoradiographic method[7,10]. It has been suggested that DA is stored in smooth endoplasmic reticulum[10,17,22].

DA metabolism in the SN

A guide to the metabolism of DA is commonly provided by measurement of its metabolites, methoxytyramine (3-MT), dihydroxyphenylacetic (DOPAC) and homovanillic (HVA) acids. This approach has also been applied to the rat SN to investigate DA metabolism in dendrites in various experimental conditions and to compare the results with those obtained in the striatum.

Korf *et al.*[24,25] observed an enhanced accumulation of DOPAC and HVA not only in the striatum but also in the SN shortly after the electrical stimulation of the middle forebrain bundle through which dopaminergic fibres run towards the striatum. The changes in the levels of nigral DA metabolites led

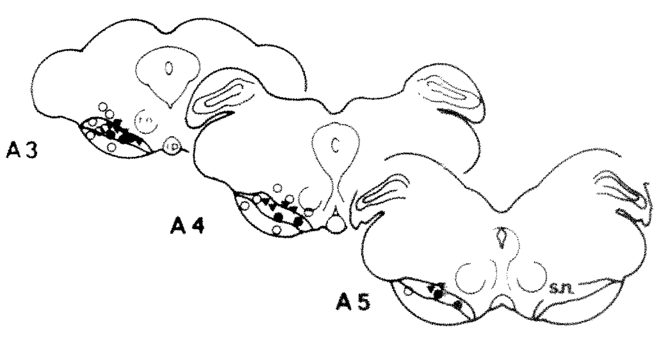


Fig. 1 *In vivo* spontaneous release of newly formed ^{3}H-DA in the cat SN. Halothane-anaesthetized cats were implanted with a push–pull cannula in the left. ^{3}H-tyrosine (50 Ci mmol^{-1}, 50 μCi ml^{-1}) was delivered continuously in the push–pull cannula (2 ml h^{-1}) and ^{3}H-DA was estimated in 10-min superfusate fractions 1 h after onset of superfusion, at a time when the spontaneous release of the ^{3}H-transmitter reached a steady state. In each case, the mean spontaneous release of ^{3}H-DA was estimated on the basis of results obtained in five successive fractions and the localization of the tip of the push–pull cannula was determined. The spontaneous release of ^{3}H-DA was much higher when the tip of the cannula was in the SN (●, >0.5 nCi per 10 min; ▲, 0.1–0.5 nCi per 10 min) than when it was in the lateral part or outside of the SN (○, <0.1 nCi per 10 min). A3, A4, A5: frontal planes of the stereotaxic atlas of Smidea and Nicmek; r.n., red nucleus; i.p., interpeduncular nucleus; s.n., substantia nigra.

0028-0836/81/070537-06$01.00

Table 1 Release of dopamine from substantia nigra slices (S) or dendrosomes (D)

Treatment			Effect	Blocked by:	Refs
(S)(1)	K^+	2.4×10^{-2}M	+	Ca^{2+}-free medium or high-Mg^{2+} medium	21
(S)(1)(2)	*d*-Amphetamine	10^{-5}M	+		37
(S)(3)	K^+	4.7×10^{-2}M	+		10
(S)(1)	*d*-Amphetamine	10^{-5}M	+		38
	GABA (K^+-evoked release)	10^{-4}M	+	Picrotoxin	38
(D)(3)	*d*-Amphetamine	10^{-4}M	+		35
(S)(1)	Substance P	10^{-5}M	+		39
(S)(1)	K^+	5.0×10^{-2}M	+	Ca^{2+}-free medium	40
	Glycine	10^{-4}M	+	Strychnine	40
(D)(1)	Substance P*	10^{-8}M	+		36
	GABA*	10^{-5}M	−		36
(S)(1)	Veratridine	3.0×10^{-6}M	+	Tetrodotoxin	33

(1), Preloaded with exogenous ^{3}H-DA; (2), endogenously synthesized ^{3}H-DA; (3), endogenous. +, Increased; −, decreased.
* Derived from estimation of tissue ^{3}H-DA levels.

them to suggest that DA was released from dendrites as a result of antidromic activation of the dopaminergic neurones.

Various drugs acutely influence the nigral DA metabolism, sometimes in a manner different from that in the striatum. For example, amphetamine reduces DOPAC formation in both structures whilst increasing 3-MT levels only in the striatum[26], whereas reserpine, which depletes striatal and nigral levels of DA[27], increases the levels of DOPAC and 3-MT in the striatum, but not in the SN[26–28]. In addition, γ-hydroxybutyrate, a general anaesthetic which enhances striatal DA levels and synthesis, induces opposite effects in the SN[29,30]. These and other[27,28,31,32] discrepancies also indicate that changes in nigral DA metabolism reflect events occurring in dendrites and not in axon collaterals.

Dendritic release of DA *in vitro*

Using slices of the rat SN, Geffen *et al.*[21] demonstrated a calcium-dependent release of ^{3}H-DA from dendrites with potassium depolarization. Similar observations were made with slices of pars compacta or pars reticulata or by measuring the release of endogenous DA[10]. A veratridine-evoked release of ^{3}H-DA sensitive to tetrodotoxin was also shown, indicating that dopaminergic dendrites have fast sodium channels[33] (Table 1). Gentle homogenization of nigral dopaminergic dendrites can produce particles which behave like synaptosomes on density gradient centrifugation[34,35]. These dendrosomes take up ^{3}H-DA which can be released by potassium depolarization[36] (Table 1). Amphetamine, a potent releasing agent of DA from nerve terminals, stimulated the release of DA from SN slices or dendrosomes[35,37,38] (Table 1). Other studies indicated that transmitters present in large amounts in the SN could also influence the dendritic release of DA (Table 1). γ-aminobutyric acid (GABA), which did not affect the spontaneous release of ^{3}H-DA newly taken up from SN slices, potentiated the potassium-induced release of the ^{3}H-amine and this effect was blocked by picrotoxin[38]. Substance P[39] and glycine[40,41] stimulated spontaneous efflux of ^{3}H-DA from rat nigral slices, the latter effect being antagonized by strychnine[40] (Table 1).

Dendritic release of DA *in vivo*

Our method for demonstrating the *in vivo* dendritic release of DA makes use of a push–pull cannula to superfuse continuously the SN of halothane-anaesthetized cats with a physiological medium containing ^{3}H-tyrosine, the precursor of ^{3}H-DA[42]. This method ensures specificity because tyrosine hydroxylase is only located in dopaminergic neurones in the SN. In fact, ^{3}H-DA represented >90% of the total content of ^{3}H-catecholamines recovered in superfusates. Furthermore, the spontaneous release of ^{3}H-DA was only substantial when the push–pull cannula was within the SN (Fig. 1). ^{3}H-DA release was reduced in the absence of calcium and markedly increased during a pulse application of potassium. Tetrodotoxin, which reduces the spontaneous release of ^{3}H-DA from nerve terminals in the caudate nucleus, induced the opposite effect when applied in the SN[42]. This effect excludes a nigral release of ^{3}H-DA from nerve terminals and suggests that some nigral afferent fibres inhibit the spontaneous dendritic release of DA. Furthermore, the inability of tetrodotoxin to reduce the spontaneous release of ^{3}H-DA from dendrites could indicate that fast sodium

Table 2 Remote induced changes in ^{3}H-DA release in both substantia nigrae(SN) of the cat

Treatment	Left SN	Right SN	Refs
d-Amphetamine in left caudate nucleus	129.1 ± 8.4*	108.6 ± 10.8	51
d-Amphetamine in left SN	—	63.7 ± 12.5*	51
α-M-*p*-T in left caudate nucleus	63.9 ± 8.3*	105.9 ± 7.3	51
α-M-*p*-T in left SN	—	119.3 ± 7.0	51
Diazepam in left SN	—	80.6 ± 5.7*	
K^+ in left SN	—	160.8 ± 40.3*	52
Light flashes in right eye	134.3 ± 6.2*	84.6 ± 5.1*	53
Electrical stimulation of right forelimb	146.5 ± 21.0*	66.7 ± 3.3*	53
Electrical stimulation of left motor cortex	199.0 ± 37.1*	Not estimated	49
Electrical stimulation of left visual cortex	157.1 ± 17.0*	Not estimated	49
Electrical stimulation of right dentate nucleus	49.3 ± 7.8*	150.2 ± 8.4*	50
Electrical stimulation of right fastigial nucleus	79.1 ± 5.9	69.6 ± 7.7*	50

Experiments were performed in cats implanted with push–pull cannulae as described in Fig. 2 legend. Changes in ^{3}H DA release are expressed as per cent of the mean spontaneous release estimated during the hour preceding the treatments. They were estimated during the first 10-min (amphetamine, potassium) or the third 10-min (other cases) fraction after the start of treatment, and are the mean of data obtained with groups of 4–7 animals compared with results from untreated cats. Data corresponding to the local effects are not indicated (—). Drug treatments: *d*-amphetamine (10^{-6} M); α-M-*p*-T (α-methylparatyrosine, 10^{-4} M); diazepam (10^{-5} M); K^+ (potassium, 30 mM).
*$P \leq 0.05$ using Student's *t*-test.

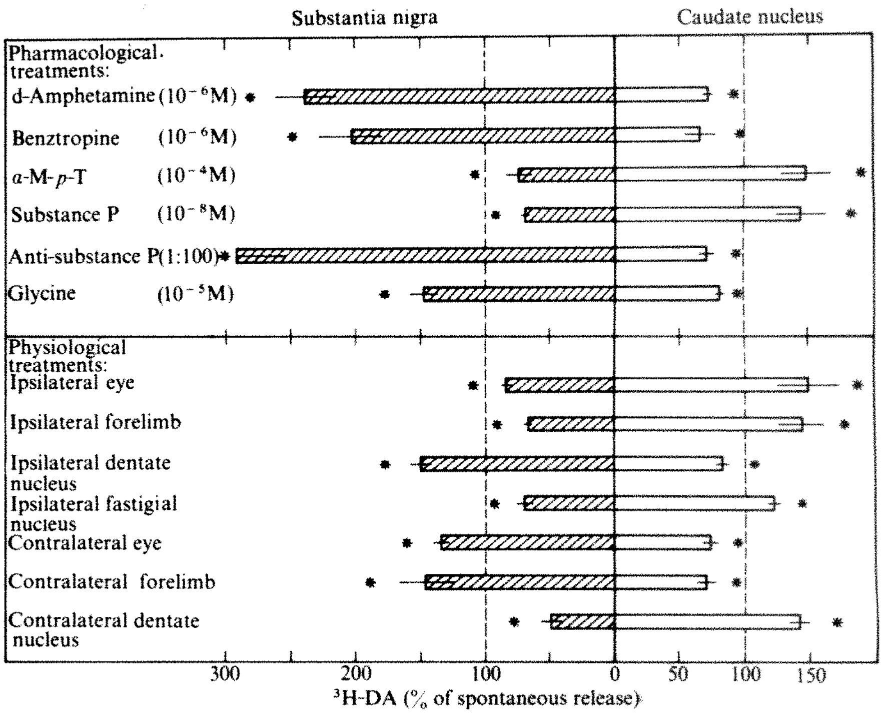


Fig. 2 Examples of treatments which induced opposite changes in the release of ^{3}H-DA from dendrites in the SN and from nerve terminals in the ipsilateral caudate nucleus. The effects of the unilateral nigral application of dopaminergic drugs, substances affecting substance P transmission, glycine or unilateral physiological stimuli on the release of newly formed ^{3}H-DA in the SN and the corresponding caudate nucleus were examined in halothane-anaesthetized cats implanted with push–pull cannulae. ^{3}H-DA was estimated in successive 10-min fractions during the continuous delivery of ^{3}H-tyrosine to each cannula. *d*-Amphetamine, benztropine, α-M-*p*-T, substance P, substance P antibody or glycine were introduced for 10–60 min into the superfusing medium delivered to the SN 3 h after the start of the experiments. The physiological stimuli comprised the delivery of light flashes applied for 10 min to the ipsilateral or contralateral eye, the electrical stimulation of the ipsilateral or contralateral forelimb (10 min), the electrical stimulation of the ipsilateral or contralateral dentate nuclei of the cerebellum (10 min) or the ipsilateral cerebellar fastigial nucleus (10 min). The results are expressed as per cent of the mean spontaneous release of ^{3}H-DA estimated during the hour preceding the treatments and correspond to changes observed during the first 10-min application of amphetamine, benztropine and glycine, or during the third 10-min fraction after the start of application of other treatments. They are the mean of data obtained with groups of 3–7 animals, compared with results from untreated cats. Statistical analysis was done using Student's *t*-test; *, $P \leq 0.05$.

channels are not involved in this process, in contrast to that observed in the caudate nucleus. In fact, dendritic spikes resistant to the neurotoxin have already been described[43–45].

Several neurotransmitters present in the SN influence the dendritic release of ^{3}H-DA. A stimulation of ^{3}H-DA release was induced by glycine (Fig. 2), acetylcholine and serotonin but substance P induced the opposite effect[46,47] (Fig. 2). Although GABA was ineffective, the GABA agonist muscimol enhanced the dendritic release of ^{3}H-DA[48] whereas the reverse was seen with diazepam, a benzodiazepine known to facilitate GABAergic transmission.

One critical step in our *in vivo* studies has been the demonstration that changes in the activity of neurones with functional connections to the SN, or involved in sensory motor processes, affected the dendritic release of DA. Thus, ^{3}H-DA release was altered by electrical stimulation of the motor or visual cortices[49] (Table 2) and of the fastigial and dentate cerebellar nuclei[50] (Table 2, Fig. 2), as well as by local pharmacological interruption, or facilitation of dopaminergic transmission in the ipsilateral caudate nucleus or in the contralateral SN[51] (Table 2, Fig. 3). Further evidence that messages originating in one SN could influence the dendritic release of DA in the contralateral structure came from the local nigral application of potassium[52] or diazepam (Table 2). Marked changes in the dendritic release of ^{3}H-DA were also induced by delivery of sensory stimuli[53] (Table 2, Fig. 2). Sometimes, the changes in ^{3}H-DA release remained for up to 1 h after treatment, suggesting the involvement of long-term regulatory processes which have still to be elucidated. These *in vivo* experiments indicate that the dendritic release of DA from the nigro-striatal dopaminergic neurones is a physiological event which must contribute to the local transfer of information.

Dendritic DA can influence the activity of ipsilateral dopaminergic neurones

Several years ago, Bunney and Aghajanian[54] suggested the presence of dopaminergic receptors on dopaminergic cell bodies or dendrites in the rat SN. They observed that the microiontophoretic application of DA or of apomorphine reduced the firing rate of the dopaminergic cells and that this effect was prevented or reversed by neuroleptics[55]. Groves *et al.*[56] reported that amphetamine (a drug which releases DA) inhibited the firing rate of pars compacta neurones after its local application and that this effect was prevented by peripheral injection of α-methylparatyrosine[57], the inhibitor of catecholamine synthesis. Because these authors also observed that the local application of haloperidol enhanced the firing rate of pars compacta neurones, they suggested that DA could be released from dendrites and was involved in a self- or lateral-inhibition of the nigral dopaminergic neurones[56]. The presence of dopaminergic autoreceptors in the rat SN has since been confirmed by binding studies with labelled spiroperidol[58–60] or apomorphine[61] after selective degeneration of the dopaminergic neurones.

To confirm the changes in the firing rate of dopaminergic cells seen in electrophysiological studies, we have measured changes in DA release from nerve terminals during pharmacological modification of the nigral dopaminergic transmission. The nigral application of DA by a push–pull cannula reduced the spontaneous release of ^{3}H-DA in the ipsilateral caudate nucleus of the cat[62] and similar results were obtained when the dendritic release of ^{3}H-DA was enhanced by the local application of amphetamine or benztropine[63] (Fig. 2). Conversely, the release of ^{3}H-DA increased in the caudate nucleus during the nigral application of α-methylparatyrosine[51] (Figs 2, 3) or of dopaminergic blockers[63], indicating that DA released from dendrites tonically inhibits the activity of the nigral dopaminergic cells. Opposite changes in the dendritic release of ^{3}H-DA and the release of the ^{3}H-transmitter from nerve terminals were also seen during modifications of glycinergic and substance P-ergic transmission in the SN (Fig. 2). The simplest hypothesis is that DA released from dendrites exerts its inhibitory effect through dopaminergic autoreceptors situated on dopaminergic cell bodies or dendrites.

We have also obtained evidence to support earlier biochemical[64,65] and electrophysiological[66] data consistent with the hypothesis of a neuronal feedback loop involved in the control of the activity of the nigro-striatal dopaminergic neurones[67]. Thus, when amphetamine was applied to the caudate nucleus, it not only increased local dopaminergic transmission, but also

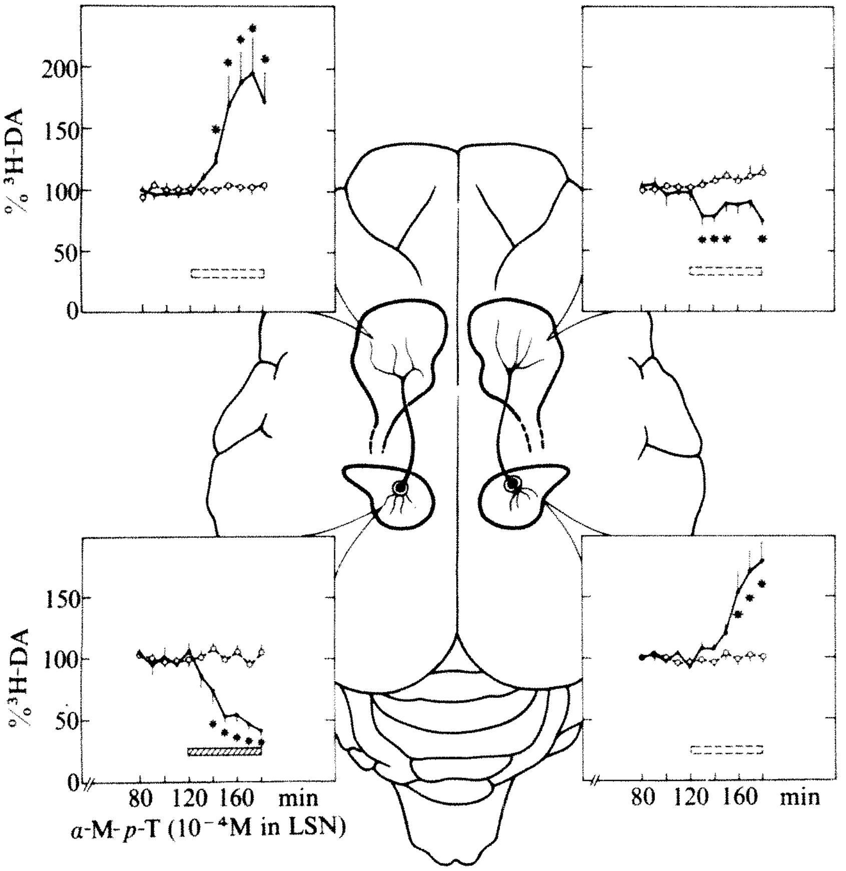


Fig. 3 Effects of α-M-p-T application into the left SN on the release of ^{3}H-DA from the two caudate nuclei and the two SN. Four push-pull cannulae were simultaneously implanted in halothane-anaesthetized cats to measure the release of ^{3}H-DA formed from ^{3}H-tyrosine (50 μCi ml^{-1}) in both SNs and both caudate nuclei. α-M-p-T was introduced for 60 min into the artificial cerebrospinal fluid superfusing the left SN (LSN) (hatched bar). In each animal and for each cannula, ^{3}H-DA in each successive fraction was expressed as per cent of an average spontaneous release calculated from the five fractions collected before treatment. Data are the mean ± s.e.m. of results obtained with five animals (●). *, $P \leq 0.05$ when compared with corresponding control values obtained in five untreated animals (○).

stimulated the dendritic release of ^{3}H-DA in the SN; α-methylparatyrosine induced the opposite effects[51] (Table 2). Therefore any changes in the dopaminergic transmission in the caudate nucleus may in turn affect the activity of the nigral dopaminergic cells through the influence of striato-nigral fibres on dopaminergic dendrites. The identity of the neurones involved in the activation or the inhibition of the dopaminergic cells has yet to be conclusively determined although substance P neurones may be responsible for activation[46,47,62,68] (Fig. 2) and GABA neurones for inhibition[69,70].

Because the modulation of the activity of dopaminergic cells by nigral afferent neurones can be mediated by their influence on the dendritic release of DA, a decreased release of DA from SN dendrites accompanied by an increased release of DA from nerve terminals in the ipsilateral caudate nucleus, and vice versa, should also be observed in physiological conditions. This pattern of responses was observed after unilateral electrical stimulation of the dentate or fastigial cerebellar nuclei[50], which send monosynaptic projections to the SN, or with the unilateral delivery of visual or somatic stimuli[53] (Fig. 2). It is not yet known whether inhibitory signals mediated by DA release from dendrites are always involved in the regulation of the activity of the dopaminergic neurones.

It will also be important to establish how DA reaches the dopaminergic autoreceptors located on parent or other dopaminergic neurones in the SN. Are there functional dendro-dendritic synapses[22] or appositions[10] as suggested by some authors or are the dendrites always separated by glial elements connected by multiple gap functions as claimed by other workers[71]? The latter features could subserve a modulatory role of DA in the SN which would explain the long-term changes in activity of the dopaminergic neurones seen in several of our *in vivo* experiments.

Dendritic DA can control nigral afferent and efferent pathways

The presence of dopaminergic receptors on nigral afferent fibres and on non-dopaminergic efferent neurones already suggests a role for dendritic DA in the presynaptic regulation of the release of other nigral neurotransmitters or in the control of the activity of non-dopaminergic cells. That postsynaptic dopaminergic receptors are associated with nigral afferent fibres is shown in the rat by the unchanged nigral DA-sensitive adenylate cyclase activity after degeneration of the nigral dopaminergic cells compared with the decrease after destruction of striato-nigral neurones[72–74].

As ^{3}H-spiroperidol binding sites not coupled to the DA-sensitive adenylate cyclase are still detectable after destruction of the dopaminergic cells, some dopaminergic receptors also seem to be located on non-dopaminergic cells in the SN[58–60]. In fact, cells in the pars reticulata have been shown to be sensitive to the microiontophoretic application of DA[75]. More recently, Ruffieux and Schulz[76] have observed that ~40% of the pars reticulata cells were excited by DA in the rat and that this effect was antagonized by fluphenazine. Furthermore, half of these cells identified by antidromic activation corresponded to the nigro-thalamic neurones. It it thus highly probable that numerous nigral non-dopaminergic neurones can be influenced by DA released from the dendrites of the dopaminergic neurones—this is also supported by recent autoradiographic studies which indicate that both apomorphine[77] and amphetamine[78] enhance glucose utilization in the pars reticulata of the SN and that the apomorphine effect is reversed by neuroleptics.

Reubi *et al.*[79] obtained the first evidence for a role of DA in the presynaptic regulation of a neurotransmitter release from nigral afferent fibres when they observed that DA and amphetamine stimulated the release of ^{3}H-GABA previously taken up in slices of the rat SN, and that these effects were blocked by neuroleptics. More recently, Van der Heyden *et al.*[80] have obtained comparable results *in vivo* although GABA release was either enhanced or decreased depending on the concentration of DA or apomorphine used. These results suggest that dopaminergic receptors coupled to the DA-sensitive adenylate cyclase are located on GABA striato-nigral fibres. In contrast, the striato-nigral substance P afferent fibres seem to lack dopaminergic receptors because the release of substance P from rat SN slices was not influenced by DA or dopaminergic agonists[81].

A series of experiments in which we investigated the mechanisms involved in the reciprocal regulation of the two nigro-striatal dopaminergic pathways demonstrated that DA released from dendrites influences the activity of non-dopaminergic nigral cells. Cats implanted with one push–pull cannula in each SN and caudate nucleus allowed us to investigate the effects of unilateral nigral facilitation or interruption of dopaminergic transmission not only on the activity of the ipsilateral dopaminergic neurones but also on the release of DA from terminals and dendrites of the contralateral dopaminergic neurones. Amphetamine, which when applied to one SN facilitated the local release of ^{3}H-DA and reduced the ^{3}H-transmitter release in the ipsilateral caudate nucleus (Fig. 2)[51,63], induced opposite affects in the contralateral SN (decreasing ^{3}H-DA release) (Table 2) and caudate nucleus (increasing ^{3}H-DA release)[51]. The unilateral nigral application of α-methylparatyrosine produced the opposite results (Fig. 3) and other experiments revealed that these contralateral effects were mediated by DA released from dendrites but not from terminals of the ipsilateral dopaminergic neurones[51]. Thus DA released from dendrites in one SN influences non-dopaminergic nigral cells involved in the regulation of the activity of the contralateral dopaminergic neurones. As no direct connections have been described between the two SN, a polysynaptic neuronal loop projecting to the contralateral SN must be involved. The nigro-thalamic neurones which can be excited by DA[76] could contribute to this neuronal loop. In fact, the contralateral effects induced by amphetamine of α-methylparatyrosine are suppressed after lesion of the thalamic massa intermedia[82]. The nigro-thalamic neurones are not the only ones to be influenced by dendritic DA—indirect evidence suggests that this is also the case for the nigral neurones projecting to the dorsal raphe. Indeed, in cats implanted with push–pull cannulae, the application of DA or α-methylparatyrosine in one SN induced changes in serotonin release in both caudate nuclei, which could reflect modifications of the activity of dorsal raphe serotoninergic neurones[83]. These results emphasize the role of dendritic DA in the control of signals leaving the SN through non-dopaminergic cells.

Unfortunately, there is as yet no firm morphological evidence for the dendro–axonic synapses suggested by biochemical data which indicate that dendritic DA is involved in the regulation of transmitter release from some nigral afferent fibres. However, Reubi and Sandri[71] have observed large special symmetrical junctions in the pars reticulata between dendrites and axons and Hattori *et al.*[17] have suggested that dendro–axonic transmission occurs in some axo–dendritic synapses in which vesicle-like structures are seen to be attached to the post-synaptic membrane. However, some of the dendro-dendritic synapses[22] or appositions[10] observed in the pars reticulata could correspond to sites by which dendritic DA influences non-dopaminergic nigral cells, as indicated by electrophysiological and biochemical studies.

Conclusions

Our knowledge concerning nigral dopaminergic transmission is now as broad as that acquired for dopaminergic transmission intervening at the level of nerve terminals in the striatum. The dendrites of the nigro-striatal dopaminergic neurones have the capacity to synthesize, store and release DA. However, the dendrites seem to differ from the nerve terminals as they contain very few vesicles and the smooth endoplasmic reticulum seems to be the storage site of the transmitter; furthermore, it is very unlikely that fast sodium channels are to be involved in the dendritic release process occurring in physiological states. Several neuronal pathways which project to the SN influence directly or indirectly, through local microcircuits, the release of DA from dendrites. As shown in Fig. 4, by acting on various types of dopaminergic receptors, only some of which are coupled to a DA-sensitive adenylate cyclase, DA released from dendrites exerts several functions. One of the most paradoxical effects of the dendritic DA is to inhibit tonically the activity of the nigro-striatal dopaminergic neurones by self- or lateral-inhibition; dopaminergic autoreceptors are involved in this phenomenon (Fig. 4:1). There is already some evidence that DA released from dendrites regulates the presynaptic release of GABA from terminals of the striato-nigral GABAergic neurones; dopaminergic receptors associated with a DA-sensitive adenylate cyclase mediate this effect (Fig. 4;2). Other dopaminergic receptors are located on some nigral neurones mainly distributed in the pars reticulata of the SN; these non-dopaminergic neurones could be interneurones involved in microcircuits (Fig. 4;3) or nigral efferent neurones (Fig. 4;3). We already know that DA exerts an excitatory effect on the nigro-thalamic neurones which contribute to a polysynaptic neuronal loop by which dendritic DA in one SN influences the activity of the contralateral nigro-striatal dopaminergic neurones. As contact between dopaminergic dendrites and glial cells[71] and capillaries[10] have been described (Fig. 4;4), the role of the dendritic DA may not be restricted to the transfer of information between neurones.

Further morphological studies at the electronmicroscopic level are required to define more precisely the relationships between dopaminergic dendrites and target cells. It also remains to be established whether or not DA released from dendrites acts at a distance from its release sites. Present biochemical techniques do not permit the monitoring of the *in vivo* release of transmitters on a time scale familiar to electrophysiologists. Therefore, the role of the dendritic release of DA during fast changes in activity of the nigral dopaminergic cells is not yet known. However, we already know that the changes in the dendritic release of DA are generally of long duration, even when they are induced by physiological treatments such as sensory stimuli. Such a phenomenon explains the prolonged modification of the activity of the dopaminergic cells observed in several situations. This time dimension is another peculiar feature of the signals mediated by dendritic DA and provides further evidence for a modulatory role of the nigro-striatal dopaminergic neurones in the transfer of information in the SN and the striatum.

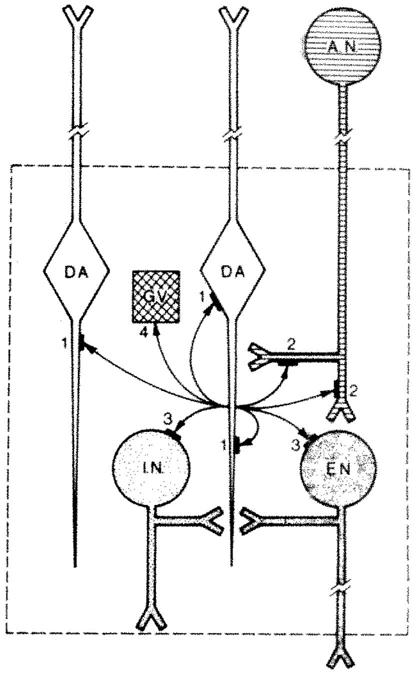


Fig. 4 Schematic representation of the various sites of action of DA in the SN: (1) dopaminergic autoreceptors; (2) dopaminergic receptors located on nerve terminals of striatonigral pathways (AN, afferent neurones); (3) dopaminergic receptors which could be located on nigral interneurones (IN) or on nigral non-dopaminergic efferent neurones (EN); (4) interactions with glial or vascular elements (GV).

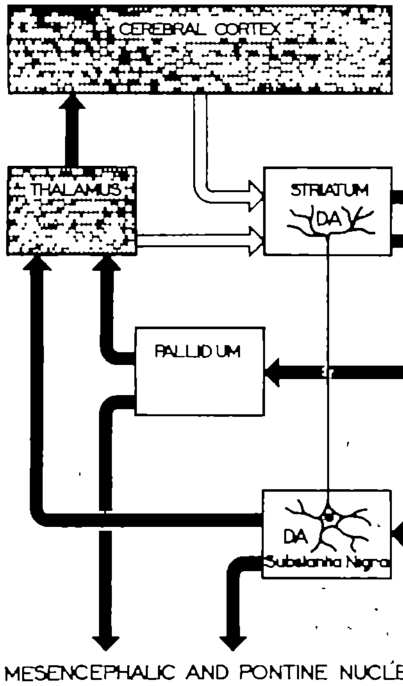


Fig. 5 Schematic diagram of some extrapyramidal circuits illustrating the role of nigro-striatal dopaminergic neurones in filtering messages travelling through the striatum and the substantia nigra. Open arrows: main inputs to the striatum; hatched arrows: main outputs from the striatum; black arrows: main outputs from the basal ganglia. DA: release of dopamine from nerve terminals and dendrites of the nigro-striatal dopaminergic neurones

Shepherd's proposal that a single neurone consists of several functional units[5] may apply to the nigro-striatal dopaminergic neurones which display functional properties both in the striatum and the SN (Fig. 5). Thus, at one level, DA released from the terminals of these nerves contributes to the fitering of messages delivered to the striatum by various afferent fibres, including the cortico- and thalamo-striatal fibres, and by influencing the activity of large populations of striatal cells it modulates the transit of signals converging on the pallidum and the pars reticulata of the SN. At another level, DA released dendritically within the SN is involved in the processing of information passing through the nigral pars reticulata in the direction of the thalamus, the subthalamic nucleus and various mesencephalic or pontine nuclei. It is thus not surprising that the nigro-striatal dopaminergic neurones have a critical role in the coordination of sensory motor processes.

One important question is to determine whether or not all central neurones share the capacity to release their transmitter from dendrites. In most cases, studies similar to those carried out on the dendrites of the nigro-striatal dopaminergic neurones are impaired by the co-existence of dendrites, nerve collaterals or nerve terminals of the same type of neurones in a given brain nucleus. However, it has already been suggested that GABA could be released from dendrites in the olfactory bulb[84], and DA released from amacrine cells in the retina[85] could originate from dendrites because this type of interneurone has no axon. Using sophisticated electrophysiological analysis, dendro-dendritic excitatory and inhibitory synapses have been particularly investigated in the olfactory bulb[5]. Furthermore, synaptic circuits through dendrites have now been found in several brain areas including the retina, the cerebral cortex, the basal ganglia, the thalamus and the suprachiasmatic and trigemenal nuclei[5]. Therefore, there is little doubt that the chemical transfer of information through dendrites is not a feature unique to the nigro-striatal dopaminergic neurones.

1 Bjorklund, A & Lindvall, O *Brain Res* **83**, 531–537 (1975)
2 Hamon, J, Pasik, P. & Pasik, T *Adv. Neurol* **14**, 149–162 (1975)
3 Kreutzberg, G W, Toth, L & Kaiya, H *Adv. Neurol* **14**, 269–282 (1975)
4 Ramon-Moliner, E *Adv Neurol* **14**, 135–148 (1975)
5 Shepherd, G. M *Scient Am* **238**, 92–102 (1978)
6. Juraska, J M, Wilson, C J & Groves, P M. *J comp Neurol* **172**, 585–600 (1977)
7 Sotelo, C *J Ultrastruct Res* **36**, 824–841 (1971)
8. Parznok, J., Hassler, R & Bak, I J *Z Zellforsch mikrosk Anat* **118**, 137–148 (1971)
9 Cuello, A C & Kelly, J S *Br J Pharmac* **59**, P 527–528 (1977)
10 Cuello, A C & Iversen, L L in *Interactions between Putative Neurotransmitters in the Brain* (eds Garattini, S, Pujol, J. F. & Samanin, R) 127–149 (Raven, New York, 1978).
11 Sotelo, C., Javoy, F., Agid, Y & Glowinski, J. *Brain Res* **58**, 269–290 (1973)
12 Hattori, T, Fibiger, H C & McGeer, P L *J. comp Neurol* **162**, 487–504 (1975)
13 Pickel, V M, Joh, T H, Field, P M, Becker, C G & Reis, D J. *J Histochem Cytochem* **23**, 1–12 (1975)
14 Hokfelt, T, Johansson, O., Fuxe, K, Goldstein, M. & Park, D *Med biol Illust* **54**, 427–453 (1976)
15 Pickel, V. M, Joh, T H & Reis, D. J. *J Histochem Cytochem* **24**, 792–806 (1976)
16 Ljungdahl, A, Hokfelt, T, Nilsson, G & Goldstein, M *Neuroscience* **3**, 945–976 (1978).
17 Hattori, T, McGeer, P L & McGeer, E G *Brain Res* **170**, 71–83 (1979)
18 Javoy, F, Hamon, M. & Glowinski, J. *Eur. J. Pharmac* **10**, 178–188 (1970)
19. Hornykiewicz, O, *Wien klin Wschr* **75**, 309–312 (1963).
20. Versteeg, D H. G, Van der Gugten, J, De Jong, W & Palkovits, M *Brain Res* **113**, 563–574 (1976)
21 Geffen, L B, Jessell, T M, Cuello, A C. & Iversen, L L *Nature* **260**, 258–260 (1976).
22 Mercer, L, Del Fiacco, M & Cuello, A C *Experientia* **35**, 101–103 (1979)
23 Wilson, C J, Groves, P M & Fifkova, E *Expl Brain Res* **30**, 161–174 (1977)
24. Korf, J, Zieleman, M & Westerink, B H C *Nature* **260**, 257–258 (1976)
25. Korf, J, Zieleman, M & Westerink, B H C *Brain Res* **120**, 184–187 (1977)
26 Westerink, B H C. *J Pharm Pharmac* **31**, 94–99 (1979)
27 Fadda, F, Argiolas, A., Stefanini, A & Gessa, G L *Life Sci* **21**, 411–418 (1977)
28 Kebabian, J W, Saavedra, J M. & Axelrod, J *J Neurochem* **28**, 795–801 (1977)
29 Hefti, F., Lienhart, R & Lichtensteiger, W *Nature* **263**, 341–343 (1976)
30 Pericic, D & Walters, J R *J. Pharm Pharmac* **28**, 527–529 (1976)
31 Wuerthele, S M, Friedle, N M & Moore, K E *J Neural Transm* **45**, 117–127 (1979)
32 Westerink, B H C & Korf, J *Eur J Pharmac* **40**, 131–136 (1976)
33 Tagerud, S. E O & Cuello, A. C *Neuroscience* **4**, 2021–2029 (1979)
34 Hefti, F. & Lichtensteiger, W *J. Neurochem.* **30**, 1217–1230 (1978)
35 Hefti, F & Lichtensteiger, W *Neurosci Lett* **10**, 65–70 (1978)
36 Silbergeld, E K & Walters, J R *Neurosci Lett* **12**, 119–126 (1979)
37 Paden, C., Wilson, C J & Groves, P M *Life Sci* **19**, 1499–1506 (1976)
38 Starr, M S, *Eur J Pharmac* **48**, 325–328 (1978)
39 Reubi, J. C., Emson, P. C., Jessell, T. M & Iversen, L L *Naunyn Schmiedebergs Archs Pharmak* **304**, 271–275 (1978)
40 Kerwin, E & Pycock, C *Br. J. Pharmac* **64**, 428 (1978).
41 Kerwin, E & Pycock, C. *Eur J Pharmac.* **54**, 93–98 (1979)
42. Nieoullon, A, Chéramy, A & Glowinski, J *Nature* **266**, 375–377 (1977)
43. Wong, R K S., Prince, D A & Basbaum, A I *Proc natn Acad Sci U S A* **76**, 986–990 (1979)
44 Fukuda, J & Kameyama, M *Brain Res.* **182**, 191–197 (1980)
45 Llinas, R & Hess, R *Proc natn Acad Sci. U S A.* **7**, 2520–2523 (1976)
46 Michelot, R, Leviel, V, Giorguieff-Chesselet, M F, Chéramy, A & Glowinski, J *Life Sci* **24**, 715–724 (1979)
47 Chéramy, A, Michelot, R, Nieoullon, A, Leviel, V & Kerdelhué, B in *Catecholamines Basic and Clinical Frontiers* Vol 2 (eds Usdin, E, Kopin, I J Barchas, J.) 1098–1100 (Pergamon, New York, 1979)
48 Chéramy, A, Nieoullon, A & Glowinski, J *Eur J Pharmac* **48**, 281–295 (1979)
49 Nieoullon, A, Chéramy, A & Glowinski, J *Brain Res* **145**, 69–84 (1978)
50 Nieoullon, A., Chéramy, A. & Glowinski, J *Brain Res* **148**, 143–152 (1978)
51 Leviel, V., Chéramy, A. & Glowinski, J *Nature* **280**, 236–239 (1979)
52 Greenfield, S, Chéramy, A, Leviel, V & Glowinski, J *Nature* **284**, 355–357 (1980)
53 Nieoullon, A, Chéramy, A & Glowinski, J *Nature* **269**, 340–342 (1977)
54. Bunney, B. S. & Aghajanian, G K. *Nature new Biol.* **245**, 123–125 (1973)
55 Aghajanian, G K & Bunney, B S *Naunyn-Schmiedebergs Archs Pharmak.* **297**, 1–7 (1977)
56 Groves, P. M, Wilson, C J, Young, S J. & Rebec, G V *Science* **190**, 522–529 (1975).
57 Groves, P M, Young, S J & Wilson, C J *Neuropharmacology* **18**, 755–762 (1976)
58 Reisine, T D, Nagy, J. I, Fibiger, H C & Yamamura, H. I *Brain Res* **169**, 209–214 (1979)
59 Quik, M, Emson, P C & Joyce, E *Brain Res* **167**, 355–365 (1979)
60 Murrin, L C, Gale, K & Kuhar, M J *Eur J Pharmac* **60**, 229–235 (1979)
61. Nagy, J L, Lee, T, Seeman, P. & Fibiger, H C *Nature* **274**, 278–281 (1978)
62 Chéramy, A, Nieoullon, A, Michelot, R & Glowinski, J *Neurosci Lett* **4**, 105–109 (1977)
63 Nieoullon, A, Chéramy, A, Leviel, V & Glowinski, J *Eur J Pharmac* **83**, 289–296 (1979)
64 Di Chiara, G., Porceddu, M L, Spano, P F & Gessa, G L *Brain Res* **130**, 374–382 (1977)
65 Garcia-Munoz, M, Nicolaou, N M, Tulloch, I F, Wright, A K. & Arbuthnott, G W. *Nature* **265**, 363–365 (1977)
66 Bunney, B S & Aghajanian, G K *Naunyn-Schmiedebergs Archs Pharmak* **304**, 255–261 (1978)
67 Carlsson, A & Lindqvist, M. *Acta pharmac tox* **20**, 140–144 (1963)
68 Chéramy, A *et al Brain Res* **155**, 404–408 (1978)
69. Bunney, B S & Aghajanian, G K in *The Basal Ganglia* (ed Yahr M P) 249–267 (Raven, New York, 1976)
70 Dray, A *Neuroscience* **4**, 1407–1439 (1979)
71 Reubi, J. C & Sandri, C *Neurosci Lett* **13**, 183–188 (1979)
72 Prémont, J *et al FEBS Lett* **68**, 99–104 (1976)
73 Spano, P F, Di Chiara, G, Tonon, G. C & Trabucchi, M *J Neurochem* **27**, 1565–1568 (1976)
74 Gale, K, Guidotti, A. & Costa, E *Science* **196**, 503–505 (1977).
75 Dray, A & Straughan, D W *J Pharm Pharmac* **28**, 400–405 (1976)
76 Ruffieux, A & Schultz, W. *Nature* **285**, 240–241 (1980)
77 Brown, L L & Wolfson, L I *Brain Res* **140**, 188–193 (1978)
78 Wechsler, L R, Savaki, H E & Sokoloff, I *J Neurochem* **32**, 15–22 (1979)
79 Reubi, J C, Iversen, L L & Jessell, T M *Nature* **268**, 652–654 (1977)
80 Van der Heyden, J A M, Venema, K & Korf, J *J Neurochem* **34**, 119–125 (1980)
81 Jessell, T M *Proc int Soc Neurochem* **6**, 220 (1977)
82 Chéramy, A., Chesselet, M F, Leviel, V & Glowinski, J (in preparation)
83 Héry, F *et al Brain Res* **193**, 143–151 (1980)
84 Jaffé, E. H & Cuello, A C (in preparation).
85 Kramer, S G *Invest Ophthalmol* **71**, 438–452 (1971)

Macroscopic quantum electrodynamic effects in a superconducting ring containing a Josephson weak link

R. J. Prance, A. P. Long, T. D. Clark, A. Widom*, J. E. Mutton, J. Sacco*, M. W. Potts, G. Megaloudis* & F. Goodall*

School of Mathematical and Physical Sciences, University of Sussex, Brighton, Sussex BN1 9QH, UK

The problem of coherent flux tunnelling across a Josephson weak link enclosed by a thick superconducting ('SQUID') ring in terms of quantum electrodynamics is considered. We show that for a low effective noise temperature ($\simeq$few K) in the SQUID ring and with the weak link capacitance $\lesssim 3\times10^{-15}$, such a process should be observable. Experimental evidence is given to justify this claim.

IN the usual treatment of a superconducting ring containing a Josephson weak link (a 'SQUID'—superconducting quantum interference device) the total flux Φ in the ring is assumed to be a classical quantity[1]. This article investigates the behaviour of such a SQUID when this assumption is no longer valid—in what we shall term the 'quantum limit', where voltage and magnetic flux measurements on the SQUID ring interfere. In this limit, we compare the predicted quantum electrodynamic behaviour with the actual experimental properties of a SQUID ring, biased at 430 MHz, in the temperature range 4.2–1.8 K.

Properties of the SQUID ring

With Φ classical, the following thermodynamic arguments macroscopically describe the properties of a SQUID ring containing one weak link. The condensed matter hamiltonian depends on Φ nonlinearly. The resulting condensed matter free energy, periodic in the flux quantum $\Phi_0=(\pi\hbar/e)$, is given by[2]

$$F(\Phi+\Phi_0, T)=F(\Phi, T) \qquad (1)$$

where the temperature of the ring is T and Φ_0 is determined by the electric charge by virtue of the gauge breaking Landau–Ginzburg electron pair wave function[3].

The free energy can be expressed in terms of the Josephson critical current $I_c(T)$ as

$$F(\Phi, T)=F(0, T)+[\Phi_0 I_c(T)/2\pi][1-\cos(2\pi\Phi/\Phi_0)] \qquad (2)$$

If an external flux Φ_x is applied to the ring (self-inductance L), then the thermal equilibrium total flux can be obtained by the minimum free energy principle

$$A(\Phi_x, T)=\min_{\Phi}\,[F(\Phi, T)+(\Phi-\Phi_x)^2/2L] \qquad (3)$$

Equations (2) and (3) imply, for equilibrium flux, that

$$\Phi+LI_c(T)\sin(2\pi\Phi/\Phi_0)=\Phi_x \qquad (4)$$

For $I_c(T)>\Phi_0/2\pi L$, equation (4) describes the hysteretic behaviour in Φ versus Φ_x of a SQUID ring[1,2]. This behaviour forms the basis of operation of the r.f.-biased, so-called 'hysteretic' SQUID magnetometer[1] and reflects the existence of local metastable free energy minima.

As a measurement device, the hysteretic SQUID depends on long lifetimes for the metastable free energy states referred to above. We now consider how quantum electrodynamic (QE) effects may considerably shorten the lifetime of such a metastable state. In QE circuits one adds both electric field energy (capacitively) and magnetic field energy (inductively) to the microscopic condensed matter hamiltonian

$$\mathcal{H}(\Phi_x)=H(\Phi)-(\hbar^2/2C)(\partial/\partial\Phi)^2+(\Phi-\Phi_x)^2/2L \qquad (5)$$

where C is the weak link capacitance and, for simplicity, we have included only the dominant LC–photon oscillator mode at an unrenormalized frequency $\omega_0^2=(1/LC)$. The stable free energy of the SQUID is now (in principle) calculated from statistical thermodynamics

$$A(\Phi_x, T)=k_BT\ln Tr\exp-(\mathcal{H}(\Phi_x)/k_BT) \qquad (6)$$

The main difference between the classical (equation (3)) and QE (equation (6)) picture is in the dynamic lifetimes of metastable states in $F(\Phi, T)$ which participate in hysteretic modes of dynamical behaviour.

Dynamical dissipative modes

Classically, the SQUID ring admittance $Y(\Phi, \zeta)$, at fixed Φ and complex frequency ζ, has been discussed by Tinkham[4]. In general, $Y(\Phi, \zeta)$ contains dissipative current fluctuation modes which obey the oscillator strength sum rule

$$\left(\frac{2}{\pi}\right)\int_0^\infty Re\,Y(\Phi, \omega+i0^+)\,d\omega=\left(\frac{1}{L_\infty}\right) \qquad (7)$$

where the kinetic inductance[5] L_∞ depends on the geometry of the ring and local electronic plasma frequencies. In the superconducting phase, the SQUID ring admittance has the Tinkham form

$$Y(\Phi, \zeta)=(\partial^2F/\partial\Phi^2)_T(i/\zeta)+1/Z(\Phi, \zeta) \qquad (8)$$

where the frequency pole is due to the supercurrent $J=(\partial F/\partial\Phi)_T$ response and the normal current weak link shunt impedance $Z(\Phi, \zeta)$ describes the dissipative part of the current response. For superconductors, equation (7) reads

$$\left(\frac{1}{L_\infty}\right)=\left(\frac{\partial J}{\partial\Phi}\right)_T+\left(\frac{2}{\pi}\right)\int_0^\infty d\omega\,Re\,Z^{-1}(\Phi, \omega+i0^+) \qquad (9)$$

with

$$(\partial J/\partial\Phi)_T=(2\pi I_c(T)/\Phi_0)\cos(2\pi\Phi/\Phi_0) \qquad (10)$$

for the free energy parameterization of equation (2).

The growth of $I_c(T)$ as T is reduced implies that dissipation current fluctuation modes in $Z(\Phi, \zeta)$ are subtracted so as to keep

---

* Present addresses: Department of Physics, Northeastern University, Boston, Massachusetts 02115 (A.W., J.S., G.M.); Rutherford Laboratories (Appleton Division), Ditton Park, Slough, Bucks SL3 9JX, UK (F.G.).

0028-0836/81/070543-07$01.00

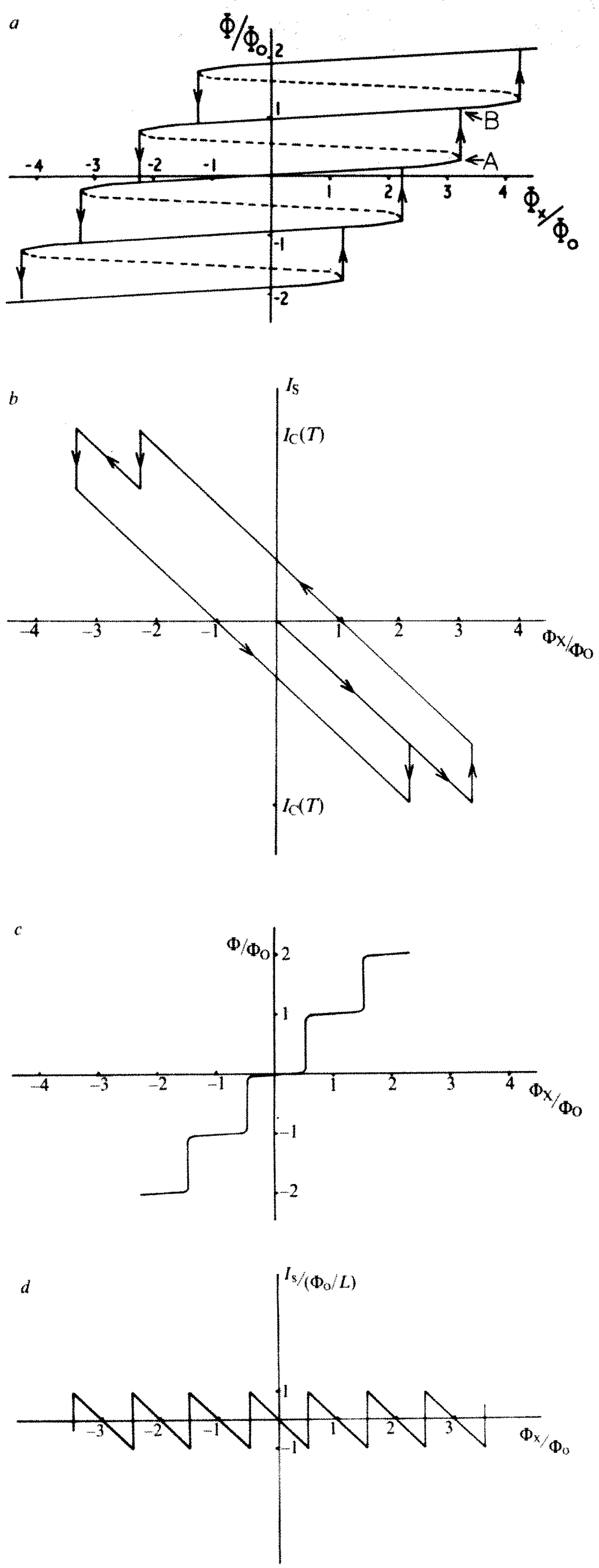


Fig. 1 *a* and *b*, Classically metastable Φ versus Φ_x and I_s versus Φ_x curves for $2\pi LI_c(T)/\Phi_0>1$; *c* and *d* Φ and I_s versus Φ_x with coherent flux tunnelling as the dominant process—again $2\pi LI_c(T)/\Phi_0>1$.

the total oscillator strength L_∞^{-1} in equation (9) constant. Tinkham has also argued, on the basis of equation (9), that the normal current dissipation in the SQUID is greatest in the region of relative maxima in $F(\Phi, T)$ and least in the region of relative minima. This allows for long lived metastability near free energy relative minima in the classical theory where Φ is treated as a classical quantity.

In QE circuits, the operator current commutators, whose average value determines the admittance, now describe, through LSZ quantum field theory, inelastic photon scattering processes[6]. Transition matrix elements are determined by an operator admittance, as discussed previously[7]. For example, an off-energy shell photon scattering amplitude, determined by $Y(\Phi, \zeta)$ considered as an operator, has the form

$$M_{\rm fi} \approx -{\rm i}\omega_{\rm fi}\int {\rm d}\Phi\cdot\Phi\chi_{\rm f}^*(\Phi)\,Y(\Phi, \omega_{\rm fi}+{\rm i}0^+)\Phi\chi_{\rm i}(\Phi) \qquad (11)$$

where the initial and final photon wave functions, $\chi_{\rm i}$ and $\chi_{\rm f}$, in the representation implicit in equation (5), describe the localization in Φ-space of the photon states. In the elastic on-energy shell limit ($\omega_{\rm fi}\to 0$) the ζ-pole in the Tinkham admittance in equation (8) dominates the lifetime of the photon states near metastable free energy minima.

The Josephson critical current $I=I_c(T)\sin(2\pi\Phi/\Phi_0)$ reflects the maximum frequency $\nu(T)$ with which coherent electron pairs can pass through a weak link contact, through $I_c(T)=2e\nu(T)$. The current matrix elements, which classically yield dissipative heating in the SQUID weak link, can, from a QE view, pass Faraday flux lines across the link. Such coherent flux tunnelling, with flux Φ_0 passing across the link, constitutes a truly macroscopic quantum phenomenon as it requires the flux Φ in the SQUID ring to be treated quantum electrodynamically. The passage of flux, in discrete units of Φ_0 across a weak link, together with an associated phase change of 2π along the link, is known generally as the phase slip process[8]. There is a critical voltage law[9,10] for coherent flux tunnelling which is directly analogous to the critical current law $\{I_c(T)=2e\nu(T)\}$ for the electron pair tunnelling. The critical voltage is $V=V_c(T)\sin(2\pi Q/q)$, where Q is the electric charge on the weak link capacitor, $V_c(T)=\Phi_0\Omega(T)$ reflects the maximum frequency $\{\Omega(T)\}$ with which flux quanta can pass across the link and $q=2e$.

Quantum electrodynamically, the main hindrance to coherent flux tunnelling arises from the small overlap of the photon wave functions $\chi_{\rm i}(\Phi)$ and $\chi_{\rm f}(\Phi)$ localized about points in Φ-space displaced by a flux quantum Φ_0. The elastic photon amplitude is determined (on-energy shell) by the supercurrent part of the SQUID admittance operator. If the noise in the capacitor charge operator $Q\to i\hbar\partial/\partial\Phi$ (which generates flux quantum displacements via $\exp({\rm i}\Phi_0 Q/\hbar)$ is gaussian, then $\Omega(T)$ and $\nu(T)$ are related by

$$\Omega(T)=\beta\pi^2\nu(T)\exp-(\Phi_0^2\langle\Delta Q^2\rangle/2\hbar^2) \qquad (12)$$

where β is a coefficient of order unity which depends on the phase slip amplitude in the flux tunnelling process.

Let us suppose that the charge fluctuation (equation (12)) is determined by an effective quantum noise temperature, T^*, defined by the 'equipartition' rule

$$\langle\Delta Q^2\rangle = Ck_{\rm B}T^* \qquad (13)$$

For a classical capacitor in thermal equilibrium $T^*=T$. Equations (12) and (13) then relate the critical voltage to the critical current by

$$V_c(T)=\beta(\pi^3/2)(\hbar/e^2)I_c(T)\exp[-\pi^2(Ck_{\rm B}T^*/2e^2)] \qquad (14)$$

For comparison the critical current for coherent pair tunnelling is given by

$$I_c(T)\simeq I_c(0)\exp[-q^2\langle\Delta\Phi^2\rangle/2\hbar^2] \qquad (15)$$

with $q=2e$ and $\langle\Delta\Phi^2\rangle\simeq k_{\rm B}TL$ by the 'equipartition' rule.

The exponential fall in $V_c(T)$ with noise temperature T^* is similar to the decay in $I_c(T)$ with increasing temperature T. With

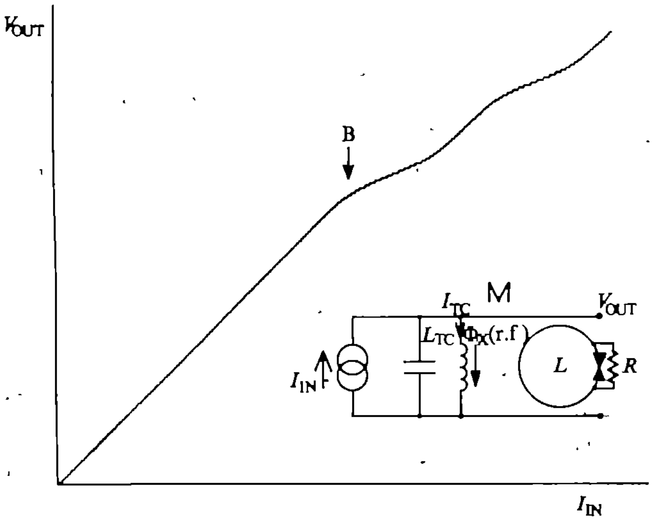


Fig. 2 V_{OUT} versus I_{IN} characteristic of a 'hysteretic' SQUID magnetometer biased at 430 MHz; $T=4.2$ K; post-detection bandwidth $=1$ Hz; $\Phi_{XDC}=n\Phi_0$(n integer); point B denotes critical current breakpoint; V_{OUT}(break)$/G=1.3$ mV, where G is the system gain.

$V_c(T)$ large compared with the total noise voltage in the SQUID, QE flux tunnelling will render the metastable free energy states (equation (4)) too short lived for the SQUID to operate in a hysteretic mode. In the other extreme, in which the SQUID noise voltage is large compared with $V_c(T)$, direct flux tunnelling will not be observed, and the SQUID ring will operate in the usual classically metastable manner. The maximum phase coherent coupling energy associated with electron pair tunnelling is $\Phi_0 I_c(T)=h\nu(T)$. Josephson effects become apparent when the weak link phase locks against thermal fluctuations, that is when $\Phi_0 I_c(T)>k_B T$. Similarly, we expect coherent flux tunnelling to be observable when the conjugate flux tunnelling phase coherent coupling energy $qV_c(T)=\hbar\Omega(T)$ becomes larger than $k_B T^*$.

Discussion

The behaviour of a classically 'hysteretic' SQUID ring (and hence that of hysteretic r.f. SQUID magnetometers[1]) can be analysed conveniently in terms of the screening supercurrent I_s circulating in the ring in response to an externally applied flux Φ_x. Figure 1*a* and *b* show, respectively, Φ plotted against Φ_x and I_s plotted against Φ_x for $I_c(T)>\Phi_0/2\pi L$. The classical metastability is apparent. In this regime, the 'fluxoid' states (A and B in Fig. 1*a*) are well defined[1]. This metastability is reduced considerably if coherent flux tunnelling becomes important. To calculate $I_s(\Phi_x)$ in this QE regime we consider the amplitude (D_n) for the SQUID ring to contain n flux quanta and look for the ground state energy $E(\Phi_x)$ of the equation

$$\{(n\Phi_0-\Phi_x)^2/2L\}D_n-(\hbar\Omega/2)(D_{n+1}+D_{n-1})=ED_n \quad (16)$$

Here, the energy stored in the ring inductor (L) is $(n\Phi_0-\Phi_x)^2/2L$, $\hbar\Omega/2$ is the flux tunnelling matrix element linking states of the ring different in flux by $\pm\Phi_0$ and, for simplicity, we take $T=0$.

Equation (16) generates macroscopic band spectra for the SQUID ring such that $E(\Phi_x+\Phi_0)=E(\Phi_x)$. As we have shown previously[11], the stable screening current is given by

$$I_s(\Phi_x)=-\mathrm{d}E(\Phi_x)/\mathrm{d}\Phi_x=\sum_{n=1}^{\infty} I_n \sin\left(\frac{2\pi n\Phi_x}{\Phi_0}\right) \quad (17)$$

Figure 1*c* and *d* show, respectively, Φ plotted against Φ_x and I_s plotted against Φ_x in the regime $I_c(T)>\Phi_0/2\pi L$ but with coherent flux tunnelling as the dominant process. In this situation all classical metastability disappears and both Φ and I_s versus Φ_x follow what we term stable, thermal curves[2].

The practical r.f.-biased SQUID magnetometer consists of a SQUID ring inductively coupled to a parallel LC tank circuit, as shown inset in Fig. 2. The coupled tank circuit–SQUID ring combination is excited at, or close to, resonance by means of an external r.f. current source (I_{IN} in Fig. 2). The tank circuit inductance is L_{TC}. The mutual inductance between this coil and the SQUID ring is M, the ring inductance is L and the weak link is represented by a supercurrent channel ($I_c(T)\sin(2\pi\Phi/\Phi_0)$, shaded double triangle) in parallel with a shunt resistor R. This parallel combination is usually referred to as the resistively shunted junction (RSJ) model[12] of a weak link. The external flux Φ_x in this system is the a.c. flux in the tank circuit inductor. In terms of this equivalent circuit we wish to determine the system response V_{OUT}, across the tank circuit, as a function of I_{IN} which we can control externally. In operation, weak coupling is maintained between the tank circuit and SQUID ring. The loaded quality ('Q') factor of the tank circuit is usually in the range 50–100.

For classically metastable behaviour (Fig. 1*a*, *b*) the SQUID ring–tank circuit combination has a linear V_{OUT}–I_{IN} characteristic up to a critical flux $\pm\Phi_{XC}$(r.f.) in the tank circuit coil. At these values of flux, and in the absence of fluctuations[13,14], almost discontinuous transitions take place between adjacent fluxoid states of the SQUID ring. The energy required to drive the SQUID ring through the hysteresis loops in Φ(r.f.), or I_s(r.f.), versus Φ_x(r.f.) is drawn from the tank circuit. This forces the SQUID ring–tank circuit into a relaxation oscillation mode[1]. With the r.f. current source amplitude $|I_{IN}|$ increasing linearly with time these oscillations lead to the well known, almost constant voltage, current steps in V_{OUT} versus I_{IN} (ref. 1). The current change along each step corresponds to the energy required to compensate for the losses in the hysteresis loops in Φ(r.f.) versus Φ_x(r.f.). The noise, either external or intrinsically generated in the SQUID ring junction 'shunt resistor', increases the slope $\mathrm{d}V_{OUT}/\mathrm{d}I_{IN}$ on the steps. In certain conditions the so-called 'step slope' α is proportional to the intrinsically generated noise flux $\langle\Delta\Phi^2_{INT}\rangle^{1/2}/\mathrm{Hz}^{1/2}$ in the ring.

Figure 2 shows a V_{OUT}–I_{IN} characteristic of an ultra-low noise SQUID magnetometer developed in our laboratory[15]. The bias frequency is 430 MHz, the temperature T of the SQUID ring and tank circuit is 4.2 K, and the post-detection bandwidth for this plot is ~1 Hz. The d.c. flux Φ_{XDC} on the SQUID ring has

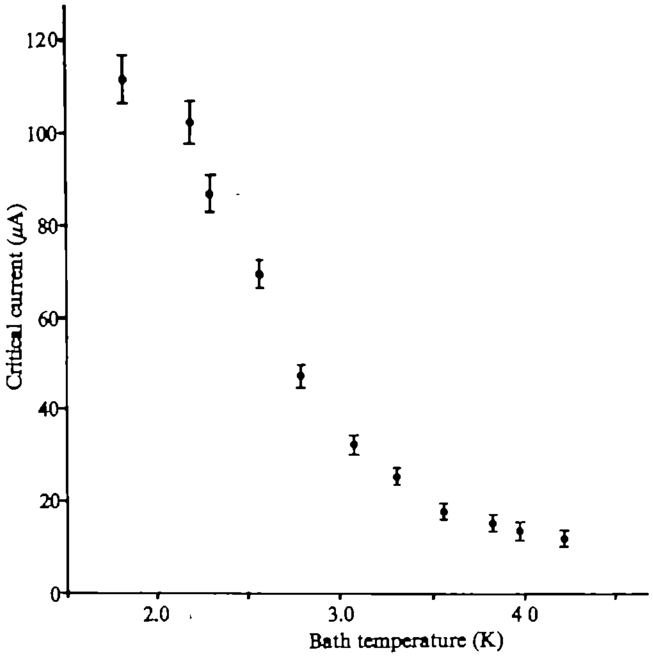


Fig. 3 Critical current $I_c(T)$ of 430 MHz magnetometer SQUID ring weak link versus temperature T of low temperature section (SQUID + tank circuit + GaAs FET preamplifier).

been set at $n\Phi_0$ (n integer). The distinctive features of this magnetometer are: (1) a liquid helium cooled GaAs FET preamplifier; (2) *in situ* (4.2 K) tuning of the GaAs FET Amp and tank circuit with GaAs varactor diodes; (3) very low noise room temperature SQUID electronics and; (4) very extensive shielding of the SQUID electronics subsystems, both at 4.2 K and at room temperature.

This magnetometer uses a two-hole 'Zimmerman' SQUID[16] consisting of a SQUID block and two screws to make the point contact weak link. The block and screws are machined from niobium, as are the lock nuts used to stabilize mechanically the point contact[17]. The point contacts are adjusted (and stabilized) at 4.2 K (ref. 15). Before the contact is made, the post (flat-ended) screw is oxidized in air at 200 °C until a light brown oxide has formed on the flat[15].

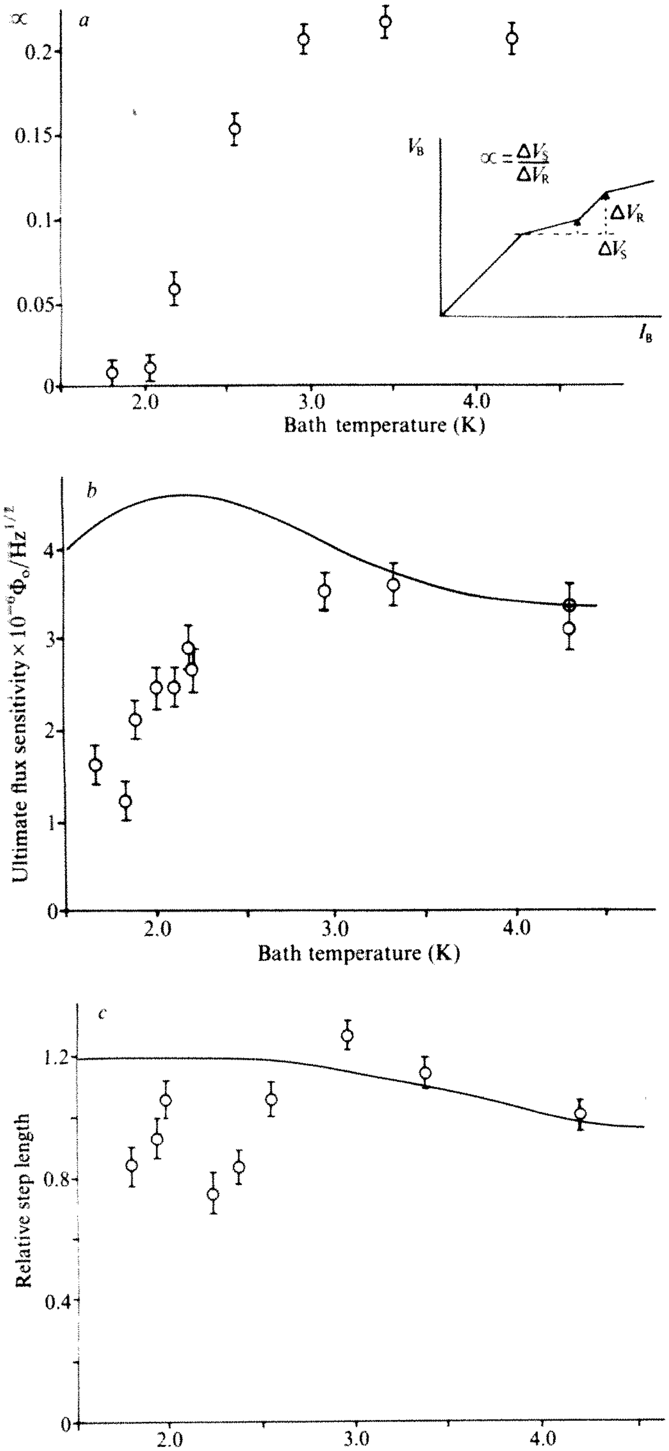


Fig. 4 *a*, Plot of 'step slope' parameter α (definition inset) versus T. *b*, Dependence of total magnetometer system flux noise on T; full line $\langle\Delta\Phi^2_{\rm NTOT}\rangle^{1/2}$ calculated from the classical theory of SQUID magnetometers. *c*, Plot of current length of first CM step versus T; full line α versus T calculated from classical theory of SQUID magnetometers.

The output of both single and two-hole SQUID magnetometers varies periodically as the input signal flux is changed, with period Φ_0. (J. E. Zimmerman, personal communication). For this periodicity the flux sensitivity (flux signal/noise ~ 1) of our 430 MHz system $\leq 4\times10^{-6}\Phi_0/\text{Hz}^{1/2}$ (conservatively) at $T = 4.2$ K, close to the value estimated from classical SQUID theory[1].

In this magnetometer the tank circuit inductance $L_{\rm TC} = 4\times10^{-8}$H, the mutual inductance (tank circuit/SQUID ring) $M = 5\times10^{-10}$H, the ring inductance $L = 5\times10^{-10}$H, $K^2Q = 0.62$ where K is the tank circuit–ring coupling coefficient, and Q(loaded) $\simeq 50$. The seven-turn tank circuit coil is constructed from either lead-plated copper or niobium foil. The GaAs FET preamplifier, with 50 Ω input impedance, is matched capacitively to the resonant tank circuit. The noise temperature $T_{\rm NA}$ of this amplifier has been measured in the Sussex and Appleton laboratories in several different ways[15]—hot/cold noise source comparisons, noise tube calibrations and direct measurement of the output noise voltage on a real-time oscilloscope. We find $T_{\rm NA}$(hot/cold) ≤ 13 K(Sussex), <10 K(Appleton); $T_{\rm NA}$(noise tube) < 10 K(Appleton) and $T_{\rm NA}$(direct) < 6 K(Sussex).

The 'break' in the slope of the $V_{\rm OUT}$–$I_{\rm IN}$ characteristic at the start (point B) of the first current step in Fig. 2 can be used to estimate $I_{\rm C}(T)$ of the SQUID ring weak link. Thus, the critical current can be estimated directly from the voltage $V_{\rm OUT}$(break) because classically[1,18]

$$V_{\rm OUT}(\text{break}) = (2\pi\nu_0 I_{\rm C}(T)L_{\rm TC}L)/M \qquad (18)$$

Alternatively, $I_{\rm C}(T)$ can be calculated from the measured change in $V_{\rm OUT}$(break) as the d.c. signal flux on the ring is changed by $\Phi_0/2$ (ref. 18). Point B in Fig. 2 is the 'break point'. In the classically metastable model of SQUID behaviour (Fig. 1*a*, *b*) no structure should be observed in the $V_{\rm OUT}$–$I_{\rm IN}$ characteristic below the breakpoint voltage $V_{\rm OUT}$(break).

We have estimated the shunt resistance R and capacitance C of the SQUID ring junction in several different ways. First, in our magnetometer there is no sign of the splitting of the current steps in $V_{\rm OUT}$–$I_{\rm IN}$ which is known, experimentally[19] and by simulation[20], to occur when $\nu_0 = \omega_0/2\pi > R/L$. There is also no obvious deterioration in the step pattern up to at least the twentieth step at $T = 4.2$ K. From these observations we conclude that $\nu_0(= 430\text{ MHz}) \ll R/L$ and $R \geqslant 10\ \Omega$ for $I_{\rm C}$(4.2 K) in the range 5–10 μA. We have also constructed a niobium 'SQUID' ring essentially identical to our usual magnetometer rings except that the post screw is electrically insulated from the SQUID block. The point contact is again adjusted at 4.2 K and stabilized mechanically with lock nuts. This arrangement enables four-terminal d.c. current–voltage measurements to be made on niobium point contact junctions prepared in the same manner as those used in our SQUID magnetometer. We find that for the post flats oxidized thermally to a light brown colour, junction resistances $\simeq 10\ \Omega$ correspond to $I_{\rm C}$ values in the range 5–10 μA at 4.2 K. This is the typical critical current range for properly adjusted r.f.-biased hysteretic SQUID magnetometers.

The RSJ model of a weak link places an upper bound on the value of C. In this model, and for constant current feed, the inclusion of a parallel capacitance across the weak link leads to a hysteretic d.c. current–voltage characteristic[12]. The current can take on one of two values at low voltages. The so-called 'McCumber' parameter $2\pi I_{\rm C}(T)CR^2/\Phi_0$ is a function of the lower and upper values of the current at zero time-averaged voltage. The lowest temperature we can reach in our cryostat is 1.8 K, at which temperature well prepared thermally oxidized point contacts with $R \simeq 10\ \Omega$ display no hysteresis. We find that

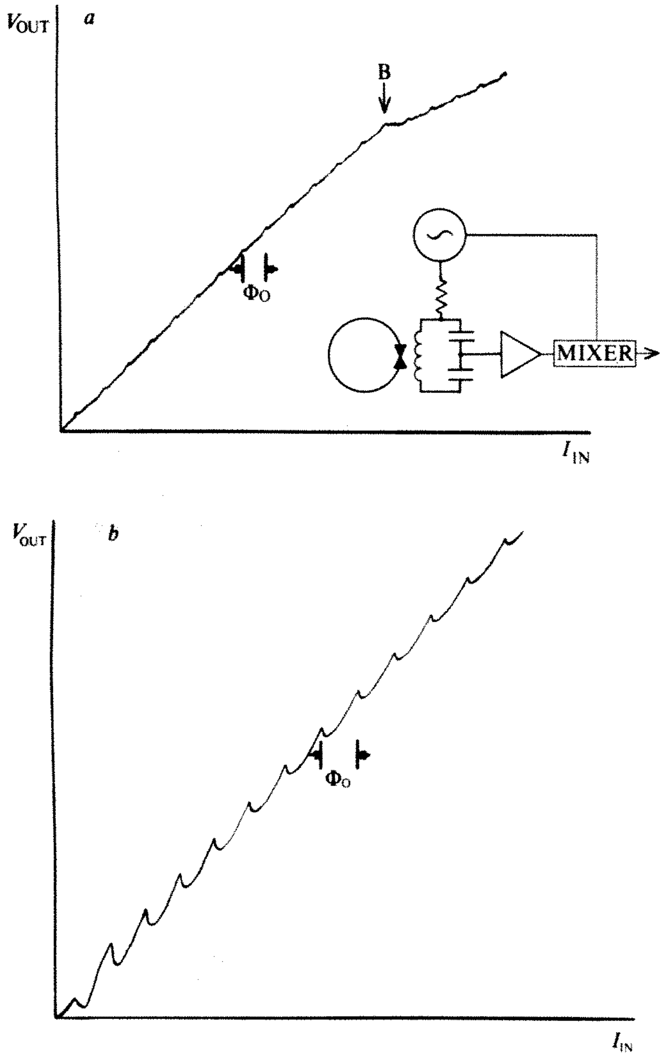


Fig. 5 *a*, Experimental V_{OUT} versus I_{IN} characteristic of the 430 MHz magnetometer of Fig. 2 at $T = 1.8$ K; post-detection bandwidth $\simeq 1$ Hz; $\Phi_{XDC} = n\Phi_0$ (n integer); point B denotes CM critical current breakpoint; $V_{OUT}(\text{break})/G = 12.3$ mV; noise in this figure and in Fig. 2 due to digitizing error on output of signal analyser used to record these curves. *b*, Theoretical V_{OUT} versus I_{IN} characteristic calculated using the coherent flux tunnelling model; $\Phi_{XDC} = n\Phi_0$(n integer).

$I_C(T)$ increases from 5–10 μA at 4.2 K to 50–100 μA at 1.8 K. This behaviour is characteristic of superconductor–normal metal–superconductor weak links[21]. For $I_C(1.8\text{ K}) \simeq 100$ μA, $R \simeq 10\ \Omega$ and taking the recorder linewidth as the maximum degree of hysteresis, we find that $C < 5 \times 10^{-15}$ F (ref. 22).

Another estimate of C can be obtained from the shape of the weak link resistance transition. For sufficiently small systems (here, the weak link) mean field fluctuations can lead to a partial suppression of the resistance of the system above the apparent critical temperature T_C (refs 23, 24). The dimensionality of the system can be obtained from the shape of the resistance transition and estimates can be made of the actual size of the fluctuating region. The thermally oxidized (light brown) point contacts studied so far appear to show two-dimensional behaviour. Treated as S/N/S junctions with a relative permitivity $\varepsilon_N \sim 1$ and current densities in the weak link $\sim 10^5$ to 10^6 A cm^{-2} (ref. 25), we estimate N-layer thicknesses ~10 Å, weak link dimensions ~1,000 Å square and C in the range 10^{-15}–10^{-16} F.

Although our UHF system behaves perfectly well as a classically metastable magnetometer at $T = 4.2$ K, changes in behaviour occur as the low temperature section (SQUID ring, tank circuit and GaAs FET Amp) is cooled from 4.2 to 1.8 K. Our main findings are:

(1) $I_C(T)$, estimated from $V_{OUT}(\text{break})$, increases rapidly as the temperature is lowered below 4.2 K. Figure 3 shows $I_C(T)$ plotted against T. $I_C(T)$ tends towards saturation at the lowest attained temperature (1.8 K).

(2) Below 3 K the 'classical metastability' (CM) current steps in V_{OUT} versus I_{IN} begin to smear out and at 1.8 K essentially all the CM steps except the first have disappeared.

(3) Between 4.2 and 1.8 K the step slope of the remaining first CM current step in V_{OUT} versus I_{IN} goes essentially to zero. A plot of this 'step slope' parameter α as a function of temperature is shown in Fig. 4*a*. The conventional definition of α is shown inset. According to the theory for CM SQUID magnetometers the intrinsic flux noise $\langle\Delta\Phi_{INT}^2\rangle^{1/2}$ is related to α by the expression[1]

$$\alpha \simeq 1.4\nu_0^{1/2}\langle\Delta\Phi_{INT}^2\rangle^{1/2}/\Phi_0\ \text{Hz}^{1/2} \qquad (19)$$

From the value of α at 1.8 K in Fig. 4*a* we calculate that $\langle\Delta\Phi_{INT}^2\rangle^{1/2} \simeq 3 \times 10^{-7}\Phi_0/\text{Hz}^{1/2}$ at 1.8 K. This is comparable with the estimate[26] of the 'quantum-limited' noise flux in the SQUID ring at finite frequency $\langle\Delta\Phi_{NO}^2\rangle^{1/2} \sim [h \times 1 \times 2L]^{1/2} \simeq 3.9 \times 10^{-7}\Phi_0/\text{Hz}^{1/2}$ for $L = 5 \times 10^{-10}$H.

(4) The total system flux noise (that is the flux noise/Hz$^{1/2}$ measured at the output of the magnetometer) decreases rapidly as the temperature is reduced below 3 K and $\simeq 1 \times 10^{-6}\Phi_0/\text{Hz}^{1/2}$ at 1.8 K. This dependence is shown in Fig. 4*b*. The full curve is the sum of the contributions from the intrinsic, preamplifier and tank circuit noise calculated using the theory for CM SQUID magnetometers[1]. Here, we take $L = M = 5 \times 10^{-10}$H, $L_{TC} = 40$ nH, $K^2Q = 0.62$, $Q \simeq 40$ and $T_{NA} \simeq 6$ K (ref. 15). Since the tank circuit is essentially power matched to the GaAs FET preamplifier we set the tank circuit noise temperature $T_{NTC} = T_{NA}$. The measurement of the system flux noise is made on the first CM SQUID current step. Both the signal and the noise are decreasing between 3 and 1.8 K.

(5) According to the theory of CM SQUID magnetometers multiple fluxoid transitions should set in when the SQUID ring damping parameter[1]

$$\gamma = \left[\frac{(2\pi I_C(T))}{\Phi_0}\right]^{1/2} RC^{1/2} \gtrsim 1 \qquad (20)$$

Taking $I_C(T) = 114$ μA at 1.8 K, $R \simeq 10\ \Omega$ and $C \leq 5 \times 10^{-15}$ F we find $\gamma = 0.4$.

In the usual interpretation of CM SQUID magnetometer behaviour[1] the length of the current step(s) should increase once multiple fluxoid transitions become possible. Figure 4*c* shows the current length of the first CM step, relative to its value at 4.2 K, as a function of temperature. The full curve is the expected CM SQUID behaviour. These data presumably eliminate the possibility of the onset of multiple fluxoid transitions providing an explanation for the observed behaviour of the SQUID magnetometer. It has been suggested (A. J. Leggett, personal communication) that heating effects in the point contact might be responsible for the decrease in step slope (α) and SQUID magnetometer signal between 3 and 1.8 K. However, theoretical studies[27] suggest that heating is not important in Nb contacts with relatively low critical currents ($\simeq 100$ μA) at temperatures $\simeq 2$ K.

(6) At temperatures $T < 3$ K weak structure appears below $V_{OUT}(\text{break})$ in the V_{OUT} versus I_{IN} characteristic. Figure 5*a* shows this characteristic taken at 1.8 K in a post-detection bandwidth of ~1 Hz. The low temperature and UHF phase sensitive detection circuits of this magnetometer are shown inset in Fig. 5*a*. The UHF phase of V_{OUT} with respect to I_{IN} has been adjusted to maximize the rounded cusp structure in the characteristic. These cusps are periodic in I_{IN} with a half period

between the first cusp and the origin. The V_{OUT} versus I_{IN} characteristic retains two CM features—the breakpoint and the remnants of the first current step following this breakpoint. In Fig. 5*a* the d.c. flux Φ_{XDC} has been adjusted to follow the upper extremum of the first current step—conventionally[1] this would correspond to $\Phi_{XDC} = n\Phi_0$ (n integer). From the measured $I_C = 114\ \mu A$ at 1.8 K the screening flux in the SQUID ring (LI_C) is $13.5\Phi_0$. No structure of any kind is observed below the breakpoint at temperatures $T > 3$ K, at least within the limits of resolution of the detection system—approximately two orders of magnitude greater than that displayed in Fig. 5*a*.

Conclusions

We believe that the above changes in the properties of the 430 MHz magnetometer are caused by the onset of QE behaviour in the SQUID ring. As we have outlined, with QE flux tunnelling dominant all classical metastability disappears and both Φ and I_S versus Φ_x follow the stable thermal curves shown in Fig. 1*c*, *d*. In terms of the classical model of SQUID the transition from metastable (Fig. 1*a*, *b*) to stable (Fig. 1*c*, *d*) behaviour corresponds to a change from r.f. 'hysteretic' to r.f. inductive 'operation' of the SQUID magnetometer. The latter, inductive mode requires that $2\pi LI_C(T)/\Phi_0 < 1$. Clearly, this is not the case, even at 4.2 K where $I_C \simeq 10\ \mu A$ and $2\pi LI_C(T)/\Phi_0 \simeq 15$. For this parameter to be <1, $I_C(T)$ must be $<0.66\ \mu A$. Good quality, thermally oxidized niobium point contact junctions with high resistances ($\gtrsim$few Ω) and low critical currents ($\lesssim$few μA) display sinusoidal current–phase (I–ϕ) relations[28]. Such a sinusoidal I–ϕ relation generates a slow sinusoidal modulation, periodic in Φ_0, on a linearly rising background in V_{OUT} versus I_{IN} (ref. 29). The rounded cusp behaviour in Fig. 5*a* would require a highly nonlinear I–ϕ relation, which seems inconsistent with the known properties of niobium point contacts.

Most of the features of Fig. 5*a* can be generated from the QE flux tunnelling model of a SQUID ring and the equivalent circuit of a SQUID magnetometer shown inset in Fig. 2. The stable QE screening current of equation (17) contains higher Josephson harmonics which affect the impedance of the SQUID ring–tank circuit in a characteristic way. We have made a computer simulation of this circuit to generate V_{OUT} in terms of I_{IN}. For a high 'Q' tank circuit (that is, $Q \simeq 50$) we assume that the dominant contribution to the output (V_{OUT}) oscillates at the input current drive frequency, that is $V_{OUT}(t) = |V_{OUT}(\omega)| \cos(\omega t + \phi_V)$, where ϕ_V is the relative phase between the output UHF voltage and the input UHF current I_{IN}. We can then calculate I_{IN} in terms of the normalized flux $B = [MI_{TC}(\omega)]/\Phi_0$ coupled into the SQUID ring. From Kirchoff's and Ohm's laws applied to the circuit in Fig. 2 we find that

$$\mathrm{Im}\, V_{OUT}(\omega) = \left(\frac{\omega_0^2}{\omega} \cdot \frac{L_{TC}}{M} \cdot \Phi\right)[B \cos \phi_{IN} - |\tilde{I}_{IN}|] \qquad (21)$$

where $\tilde{I}_{IN} = (M/\Phi_0)I_{IN}$ and ϕ_{IN} is the phase of the input current I_{IN}.

Figure 5*b* shows the numerical solution of equation (21) for $I_{TC} = 4 \times 10^{-8}$H, $L = M = 5 \times 10^{-10}$H, and $\Phi_{XDC} = n\Phi_0$ (n integer). In this calculation R (taken $= 2\Omega$) represents dissipation in the SQUID ring–tank circuit system. Dissipation is built into the equivalent circuit for this system (Fig. 5*a*, inset) as the GaAs FET preamplifier, with a real 50 Ω input impedance, is power matched to the tank circuit. There is also a small amount of dissipation in the tank circuit itself. The solution consists of a linear ramp with cusps periodic in Φ_0. The first cusp occurs at a half period from the origin. Each cusp in V_{OUT} versus I_{IN} corresponds to a pair of jump discontinuities (one each side of the origin) in the stable UHF screening current curve $I_S(\Phi_x)$ versus Φ_x (Fig. 1*d*). There is no breakpoint in the V_{OUT} versus I_{IN} characteristic because, by definition, in this model flux periodically tunnels into (or out of) the ring to prevent the screening current building up to $I_C(T)$ (see Fig. 1*d*). However, the similarity between the experimental characteristic below the breakpoint in Fig. 5*a* and the theoretical QE flux tunnelling characteristic in Fig. 5*b* is obvious.

The relatively sharp transition from CM to QE SQUID behaviour (Figs 4*a*, *b* and 5*a*) can be explained in terms of equation (14) for the coherent flux tunnelling critical voltage $V_C(T)$ and Fig. 3 for $I_C(T)$ versus T. In this magnetometer the preamplifier (and hence the tank circuit) noise temperature T_{NA} is so low that only the intrinsic flux noise contribution is important[15], at least within the range of validity of the RSJ model. We consider that the shunt resistance R in the RSJ model is, in fact, a non-equilibrium manifestation of classical metastability[4]. As flux tunnelling begins to dominate so the real phase or flux slip processes[8] in the SQUID ring junction become progressively weaker until the classical shunt resistance in the RSJ model disappears. The classical intrinsic flux noise $\langle\Delta\Phi_{INT}^2\rangle^{1/2}$ will also then disappear provided the effective noise voltage coupled into the SQUID ring from other noise sources is less than $V_C(T)$. The only remaining source of noise is the GaAs FET amplifier. We estimate that the r.m.s. noise voltage $\langle V_{NAS}^2\rangle^{1/2}$ coupled into the SQUID ring from this amplifier through the tank circuit $\simeq 0.5\ \mu V$, for $\nu_0 = 430$ MHz, $Q \simeq 50$ and a seven-turn tank circuit coil. If we take $T^* = 2$ K (the approximate temperature T of the SQUID ring) in equation (14), $C \simeq 3 \times 10^{-15}$F, $I_C(1.8\ K) = 114\ \mu A$ and $\beta \simeq 1$ we calculate $V_C(T) = 0.9\ \mu V$. To a first approximation we expect the SQUID ring to phase-lock for coherent flux tunnelling when $\langle V_{NAS}^2\rangle^{1/2} < V_C(T)\rangle$. For $C = 3 \times 10^{-15}$F and $T^* = T$, $\langle V_{NAS}^2\rangle^{1/2} = V_C(T)$ when $T = 2.6$ K and $I_C(T) \simeq 84\ \mu A$. In the flux tunnelling model the sudden change in SQUID behaviour arises simply because of the very rapid rise in $I_C(T)$ between 3 and 2 K.

The critical voltage is defined as $V_C(T) = \Phi_0\Omega(T)$ so for $V_C(1.8\ K) = 0.9\ \mu V$ the coherent flux tunnelling frequency $\Omega(1.8\ K) \simeq 430$ MHz. The noise voltage $\langle V_{NAS}^2\rangle^{1/2}$ in the SQUID ring due to the GaAs FET Amplifier is a time-averaged quantity. There will be mean, occasionally, noise voltage excursions large compared with $\langle V_C(T)$. This may provide an explanation for the remnants of classical behaviour in V_{OUT}–I_{IN} at 1.8 K. This characteristic is a time-average in a 1-Hz bandwidth over a very large number ($\sim 10^{10}$) of sweeps in $I_S\{\Phi_x(UHF)\}$ versus $\Phi_x(UHF)$. Given our estimate of $\sim 0.5\ V_C(1.8\ K)$ for the noise voltage, we expect noise fluctuations to be large enough on some sweeps to destroy the phase coherence of the flux tunnelling process. By definition, $V_{OUT}(CM)$ must be phase-incoherent with respect to $V_{OUT}(QE)$ so in the flux tunnelling model $(V_{OUT}$–$I_{IN})_{observed}$ will be a superposition of $(V_{OUT}$–$I_{IN})_{CM}$ and $(V_{OUT}$–$I_{IN})_{QE}$. Each current step in $(V_{OUT}$–$I_{IN})_{CM}$ requires four classically metastable fluxoid transitions (two hysteresis loops for $\Phi_{XDC} = n\Phi_0$). If QE flux tunnelling dominates over classical thermally activated fluxoid transitions, successive steps in V_{OUT} versus I_{IN} will become qualitatively more and more improbable. Thus, we would expect higher order steps in V_{OUT} versus I_{IN} to be very weak and this is what we observe experimentally. From this viewpoint the magnetometer signal will be reduced either because very few full hysteresis loops are traversed per second or because QE flux tunnelling processes narrow the width of these hysteresis loops dramatically. The observed reduction in noise in the magnetometer between 3 and 1.8 K will, in this model, be due to the disappearance of the intrinsic flux noise $\langle\Delta\Phi_{INT^2}\rangle^{1/2}$. The residual flux noise, which is derived from the preamplifier, is in this magnetometer[15], comparable with $[h \times 1 \times L]^{1/2}(\mathrm{Hz})^{-1/2}$.

We conclude, therefore, that if the noise injected into a SQUID ring from external sources is reduced sufficiently so that $T^* \simeq$ few K, QE behaviour will be observed at $T = 2$ K provided the SQUID weak link capacitance $C \leqslant 3 \times 10^{-15}$F.

We thank Professor A. J. Leggett for interesting discussions on the subject of macroscopic quantum effects in superconductors and to Professor D. F. Brewer for his invaluable support, also C. Trubridge, M. Moore and R. Askew for technical assistance.

Received 2 May; accepted 3 December 1980.

1. Jackel, L. D. & Buhrman, R. A. *J. Low Temp. Phys.* **19,** 201 (1975).
2. Grassie, A. D. C. *The Superconducting State,* 123 (Sussex University Press, 1975).
3. Tinkham, M. *Introduction to Superconductivity* Ch. 4 (McGraw-Hill, New York, 1975).
4. Tinkham, M. *Introduction to Superconductivity,* 195–196 (McGraw-Hill, New York, 1975).
5. Grassie, A. D. C. *The Superconducting State,* 121 (Sussex University Press, 1975).
6. Widom, A. *J. Low Temp. Phys.* **37,** 449 (1979).
7. Widom, A. *Phys. Rev.* B **21,** 5166 (1980).
8. Rieger, T. J., Scalapino, D. J. & Mercereau, J. E. *Phys. Rev.* B **6,** 1734 (1972).
9. Widom, A. & Clark, T. D. *Nuovo Cimento Lett.* **28,** 186 (1980).
10. Blatt, J. M. *Prog. theor. Phys.* **26,** 761 (1961).
11. Widom, A., Clark, T. D. & Megaloudis, G. *Phys. Lett.* **76A,** 163 (1980).
12. McCumber, D. E. *J. appl. Phys.* **39,** 3113 (1968).
13. Kurkijärvi, J. *Phys. Rev.* B **6,** 832 (1972).
14. Kurkijärvi, J. *J. appl. Phys.* **44,** 3729 (1973).
15. Long, A. P., Clark, T. D. & Prance, R. J. *Rev. scient. Instrum.* **51,** 8 (1980).
16. Zimmerman, J. E. & Frederick, N. V. *Appl. Phys. Lett.* **19,** 16 (1971).
17. Buhrman, R. A. *et al. J. appl. Phys.* **45,** 4045 (1974).
18. Lounasmaa, O. V. *Experimental Principles and Methods Below 1 K,* 156–158 (Academic, New York, 1974).
19. Clark, T. D. & Jackel, L. D. *Rev. scient. Instrum.* **46,** 1249 (1975).
20. Simmonds, M. B. & Parker, W. M. *J. appl. Phys.* **42,** 38 (1971).
21. Clarke, J. *Proc. R. Soc.* A **308,** 447 (1969).
22. Long, A. P., Prance, R. J., Clark, T. D., Hughes, J. & Widom, A. (in preparation).
23. Kadanoff, L. D. & Laramore, G. *Phys. Rev.* **175,** 579 (1968).
24. Clark, T. D. & Tilley, D. R. *Phys. Lett.* **29A,** 514 (1969).
25. Lickarev, K. K. *Rev. Mod. Phys.* **51,** 126 (1979).
26. Buhrman, R. A. *Proc. I.C. SQUID* (eds Hahlbohm, H. D. & Lübbig, H.) (de Gruyter, Berlin, 1976).
27. Tinkham, M., Octavio, M. & Skocpol, W. J. *J. appl. Phys.* **48,** 1311 (1977).
28. Jackel, L. D., Warlaumont, J. M., Clark, T. D., Brown, J. C. & Buhrman, R. A. *Appl. Phys. Lett.* **28,** 353 (1976).
29. Hansma, P. K. *J. appl. Phys.* **44,** 191 (1973).

Monoclonal antibodies distinguish identifiable neurones in the leech

Birgit Zipser & Ronald McKay
Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, New York 11724

Monoclonal antibodies were isolated by screening 475 hybridomas obtained from mice immunized with whole leech nerve cords. The majority (about 300) reacted with leech nervous tissue, but only about 40 made antibodies that identified single kinds or small sets of cells. Twenty of the antibodies which react with specific neurones were studied in greater detail and are described here. They include antibodies against identified sensory neurones and motor neurones as well as against numerous unidentified cells.

THE ability of a nervous system to generate coherent behaviour is dependent on a vast number of precise connections between neurones. Virtually nothing is known about the molecular mechanisms which generate these connections, but all tenable hypotheses ultimately postulate the existence of molecules, differing in kind or quantity from cell to cell, which mediate the necessary recognition. The problem of neural specificity becomes apparent when one considers that neurones typically receive inputs from and send outputs to whole sets of other cells each of which in turn has its own extremely complex characteristic connection pattern. Connections are established during development but the degree to which marker molecules may be present in adults is unknown, although regeneration studies in adults[1] imply long-term persistence of at least some molecular specificity. It has long been known that there are many chemical differences between neurones, involving such functions as transmitter synthesis. Recent studies on the identification and localization of neuronal peptides[2] have greatly widened the scope of detectable molecular variation among nerve cells. The relationship of this chemical variation to the mechanism of connection is still unknown. Progress in this area will require the development of ways of relating chemical specificity to unique connectivity. Because, by hypothesis, the markers of interest will differ in each neurone with different connections, these molecules must somehow be identified in structured material where the same neurone can be easily identified in repeated experiments. Such a simple system is required not only for the identification of specific neuronal markers but also for analysing the rules which establish neuronal connections.

As a step towards this goal, we have investigated whether individual cells and sub-networks of the relatively simple leech nervous system could be distinguished by specific antibodies. As the putative antigens could not, in principle, be purified *a priori,* we used the technique of hybridoma formation[3], which allows the isolation of lymphocyte clones secreting antibodies specific for individual molecules even though a complex mixture of antigens was used for immunization. Monoclonal antibodies were obtained from lymphocytes of mice immunized with the entire isolated nervous system of the leech and screened on intact ganglia. The results of our study, reported here, give a clear answer to the question of the existence of chemical markers in individual cells and sub-networks. These specific markers are surprisingly abundant, and extrapolation from our initial sample makes it feasible that every cell has one or more chemical markers shared by only small subsets of neurones. Although we know little about the exact cellular locations of these markers, many are present in all parts of the cell, including long axonal projections and neural terminals. Indeed, the situation seems quite analogous to a colour-coded electric cable

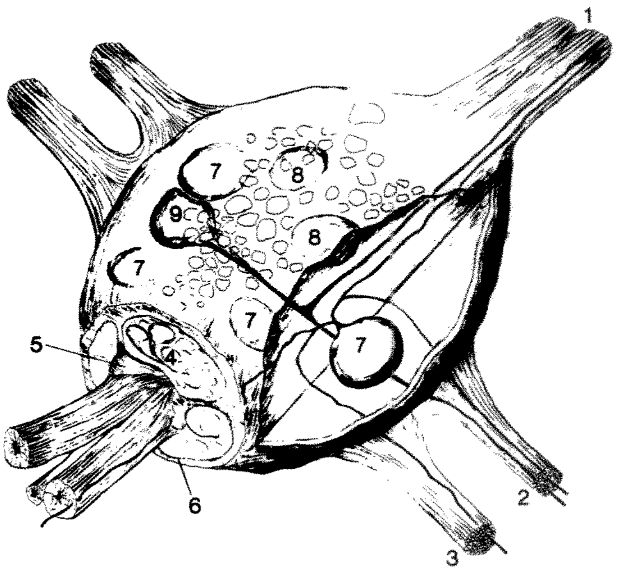


Fig. 1 Diagram of a midbody ganglion. (1) Connective, (2) anterior root, (3) posterior root, (4) neuronal cell bodies in glial packages, (5) the beginning of the neuropile where synapses occur, (6) capsule, (7) two pairs of bilaterally symmetrical pressure cells, (8) the pair of large Retzius cells are shown in the background of several hundred neuronal cell bodies, (9) one of the two lateral penile evertor cells (the other has been dissected away). Each cell body has a characteristic location and number of axons. Note that the penile evertor motor neurone projects into contralateral roots and the sensory pressure cell has ipsilateral projections.

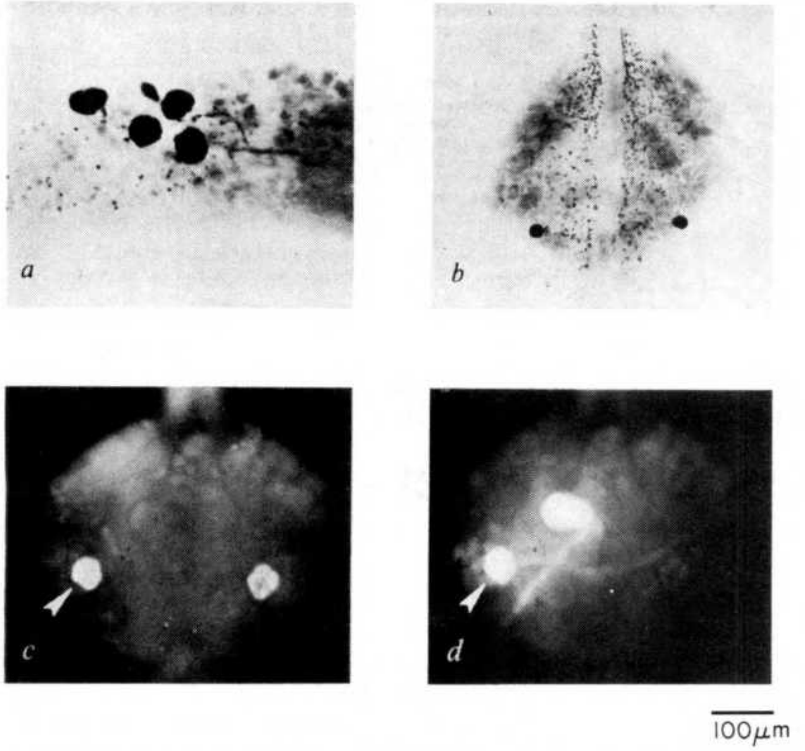


Fig. 2 Specific staining with Lan3-1. The monoclonal antibody was visualized immunocytochemically using peroxidase-conjugated (*a*, *b*) or rhodamine-conjugated (*c*) second antibody. *a* Shows cell bodies in the right supraoesophageal ganglion stained with Lan3-1. *b* Shows a bilaterally symmetrical pair of reactive cell bodies in a typical midbody ganglion. In addition to the two deeply stained cell bodies the neuropile contains a large number of stained beaded processes (varicosities) which extend into one or more axons in the connective. *c* Shows two larger cell bodies which reproducibly occur in the fifth and sixth ganglia stained with a rhodamine-conjugated second antibody. The two smaller cell bodies are also present but lie in a different focal plane. *d* Shows the same ganglion as *c* but viewed under FITC optics which reveals the presence of microinjected fluorescent dye lucifer yellow. The left lateral PE cell (marked by an arrow) was identified by a unique synaptic relationship to the morphologically and physiologically identifiable rostral penile evertor cell. The other cell labelled by lucifer yellow as a control is the Retzius cell.

containing many wires, where each wire has its own unique molecule (dye) to facilitate proper recognition and connection at terminals. The existence of specific markers is far from sufficient to explain the mechanism of neural connection specificity, but it does give us a handle by which the problem can be attacked at a previously inaccessible level. In addition to their usefulness in elucidating the mechanism of connectivity, these monoclonal antibodies will be of value in broadening our understanding of leech neurobiology, for they clearly identify whole systems of neurones together with their axonal patterns. In this article we present a survey of our initial results, together with detailed data on some of the monoclonal antibodies isolated.

The leech nervous system

The leech nervous system consists of a ganglionic chain which can be dissected out of the organism virtually intact and still electrophysiologically active. This chain has 34 copies of a basic 400-neurone ganglion in which most neurones are present as bilateral pairs. Each ganglion, in addition to containing unipolar nerve cell bodies on its outer surface, has a central neuropile where all the synaptic connections are made. The individual ganglia communicate with each other by means of a fibre tract called the 'connective' which contains thousands of fine nerve fibres. Input and output linking the ganglia to sensory, motor and secretory organs in the leech body pass through four fibre tracts called roots. A typical leech neurone has a spherical body of 10–90-μm diameter from which a stalk-like process runs a short distance into the central neuropile. Here it ramifies extensively, making many synaptic connections which are often associated with bulb-like varicosities of up to a few micrometres diameter. Many cell bodies also send processes through the connective to adjacent ganglia and to the periphery through one or more of the roots. The basic nervous system is modified by the addition of extra neurones to ganglia which have specialized functions such as reproduction, or by the fusion of ganglia to create the larger head and tail nuclei. In addition, there is a structure near the head which is of entirely different embryological origin[4]—the supraoesophageal ganglion, with many neurones some of which apparently have neurosecretory function[5]. Individual leech neurones can be reliably recognized by various features such as position, cell body diameter, electrical activity and axoneuritic pattern[1,6–8]. Many neurones, such as the sensory cells responding to pressure, touch and noxious mechanical stimulation, or the motor neurones innervating the well defined simple body wall musculature, can be visually identified through a dissecting microscope. Not only do the nerve cells occupy unique locations in the ganglia but the long axons they send into the connective occupy characteristic positions in its tightly packed bundles of fibres. A diagram illustrating a leech ganglion is given in Fig. 1. There are about 15 times more axons in the connective[9] than there are neurones in a typical midbody ganglion; the source of this large number of fibres is, in part, the cells in individual ganglia that send axons long distances down the connective, often all the way to the head and tail.

Generation and screening of antibodies

To obtain monoclonal antibodies, mice were immunized with homogenized material made from leech nerve cords (*Haemopis marmorata*) containing primarily neurones and glia, but with some muscle and connective tissue cells as well. Cells from the spleens of immunized mice were fused with myeloma cells and plated out at very low density in microtitre plates to obtain single clones. A total of three separate fusions yielded about 475 cell lines which supplied the hybridoma material for the work described here.

Previous work on enkephalin localization in the leech[10] had shown that a single neurone containing a unique antigen can be visualized in the whole, unsectioned ganglia by immunocytochemical staining techniques. These procedures served as a starting point for developing the immunocytochemistry necessary to visualize the distribution of antigens recognized by the monoclonal antibodies described here. Hybridoma growth medium was screened for the presence of interesting antibodies on segments of nerve cord containing several ganglia. The cytological staining procedure, described in detail in Table 1, legend involved dissecting out an entire ganglionic chain, mechanically breaking the capsule surrounding each ganglion, fixing the material and cutting it into segments of three or four ganglia, which were then reacted with antibody. After the appropriate washing steps this material was incubated with rabbit anti-mouse antisera to which either a fluorescent label or the enzyme horseradish peroxidase (HRP) had been conjugated. Of the 475 individual hybridomas screened, 3 came from a small preliminary test, 43 from one major fusion and 429

Table 1 Summary of staining patterns of 20 monoclonal antibodies staining specific features of leech nervous system

	No. of ganglia tested	Neurones in standard midbody ganglion	Extra neurones in specialized ganglia	Degree of generalized neuronal labelling	Processes in neuropile and fibretract
Lan3-1*	>300	2	+	0	+
2*	67	4	+	+	+
3*	36	2	+	0	0
4	30	2	+	+ +	+
5*	58	12	NT	+	+
6*	120	<30	+	+	+
Lan2-1	48	< 30	+	+	+
Lan3-7	6	<10	NT	0	0
8*	72	Complex	Complex	+ + +	S-cell axon
Lan1-1*	15	<30	NT	+	+
Lan2-2	10	<30	NT	+	+
Lan2-3*	10	<30	NT	+	+
Lan3-9*	>100	0	+	0	+
10	94	0	+	0	+
11*	77	0	+	0	+
12*	42	0	+	0	+
13*	46	0	0	0	+
Lan2-3	36	0	0	+	+
Lan2-4*	6	0	NT	0	+
Lan2-5	6	0	NT	0	+

Generation of monoclonal antibodies: Female C3H mice (Jackson Laboratories) were immunized by intraperitoneal injection at various times with the entire nerve cord of four leeches. Four days before fusion mice were immunized by intravenous injection of the SDS-soluble material from four leech nerve cords. The fusion procedure was essentially that of Kohler and Milstein[3]. The two initial fusions were performed with the SP-2 myeloma cell line[15]. Antibodies from these fusions are numbered Lan1 and Lan2. The third fusion was carried out with the P3-X63 8Ag cell line and the derived clones numbered Lan3. The hybrid cells were plated into 1,500 microwells and grown in hypoxanthine–aminopterin–thymidine medium. The hybridoma supernatant could be tested directly or as a fluid derived from an ascites tumour. Of the 41 lines identified as staining selectively on the first screen the great majority were successfully grown up and frozen so that they could be recovered up to 5 months later. The hybrids with P3-X64 as a parent are best stored at −70 °C, not in liquid nitrogen. Of the 20 lines shown 13 were recloned in soft agar and shown to be monoclonal. All lines that were recloned gave the identical staining pattern to the original hybridoma supernatant. The cloning in soft agar gave >95% positive clones on the first round as a result of the high dilution of the initial plating. These two observations together with the low frequency of specific antibody-labelling pattern (5–10%) suggested that this is a good strategy for generating antibodies of interest. Staining of leech nerve cord: The leech nerve cords were dissected from alcohol-anaesthetized leeches stored at 10 °C and fixed in 4% paraformaldehyde, 0.1 M phosphate buffer, *p*H 7.4, for 30 min at room temperature for the antibodies Lan2-3 and Lan3-7. All the other antibodies shown were bound to ganglia fixed with Bouin's fixative for 4 h at room temperature. After washing in PBS (0.05 M phosphate buffer *p*H 7.4, 0.9% sodium chloride) the capsules were cut. Hybridoma supernatant containing 0.2% saponin was incubated with the fixed ganglia overnight. Before and after addition of the second antibody the ganglia were washed in three 10-min changes of PBS, 0.2% saponin. The second antibody was either HRP-conjugated or rhodamine-conjugated rabbit anti-mouse IgG (Cappel) diluted 1:30. The peroxidase substrate used was 3′,3-diaminobenzidine. In the case of the antibodies Lan3-10 and Lan3-6 protein A was used as a link in a modification of the peroxidase anti-peroxidase (PAP) procedure[16]. Protein A was bound to the monoclonal antibody for 1 h at 37 °C, after washing in PBS and rabbit PAP (Cappel) was added at a dilution of 1:200 for 30 min. The lucifer yellow dye used in microinjection was viewed under fluorescein optics as described previously[13]. The electrophysiological characteristics of the cells controlling penile eversion have been described in detail previously. Whole mounts of leech nerve cord stained with both rhodamine and peroxidase methods were analysed by microphotometry on the Leitz microscope attached microdensitometer. The cells stained by Lan3-1, Lan3-2 and Lan3-9 both gave approximately the same 5–10-fold signal over background in the whole mounts. Thin sections (1 μm) would be expected to reduce the background noise considerably. Thin sections of the peroxidase stained supraoesophageal cell bodies stained by Lan3-9 gave a signal 50-fold over the absorbance of adjoining tissue.

* Recloned in soft agar.

from another. Of these, 64% secreted antibodies that bound to the leech nervous system. Most showed no obvious preference for subsets of neurones but reacted with general nerve cell components. A few bound to interesting subcellular or extracellular structures, for example, to antigens specific for the axon hillock of leech neurones (Lan3-15). Two antibodies recognized muscle antigens and stained the large muscle fibre present in each ganglion. Of central concern to us here is the group 41, which reacted with restricted subsets of neurones or fibres.

Antigen distribution

Of the 41 antibodies that stained restricted sets of neurones, about half have been studied in some detail and are reported here. Table 1 summarizes our findings about each of these monoclonal antibodies. The data for Table 1 are based on the immunocytochemical staining patterns of each antibody. An example is illustrated in Fig. 2 showing Lan3-1, a monoclonal antibody which binds to a pair of small bilateral neurones in a midbody ganglion (Fig. 2*b*). The right and left cell bodies are located at homologous positions on the dorsal surface in each of the 21 midbody ganglia as shown diagrammatically in a schematic representation of the leech nerve cord (Fig. 3). In addition to labelling these cell bodies, the antibody binds to clouds of varicosities in the central ganglionic neuropiles as well as to specific axons in the connective, the fibre tract that links all ganglia in the 34-ganglion nerve cord. Because each monoclonal antibody seems to have its own optimal histochemical procedures, to obtain good staining patterns, it is necessary to test a spectrum of fixatives and secondary antibodies. For example, with Lan3-1, the small pair of cell bodies appeared during the initial screen using 4% paraformaldehyde fixation, but the fine axons and varicosities were only observed after switching to Bouin's fixation. The HRP-conjugated rabbit anti-mouse IgG heavy and light chains remained a good secondary antibody but peroxidase-conjugated μ chain-specific rabbit anti-mouse IgM

LIBRARY
COLLEGE
Upper Circular Road
CALCUTTA

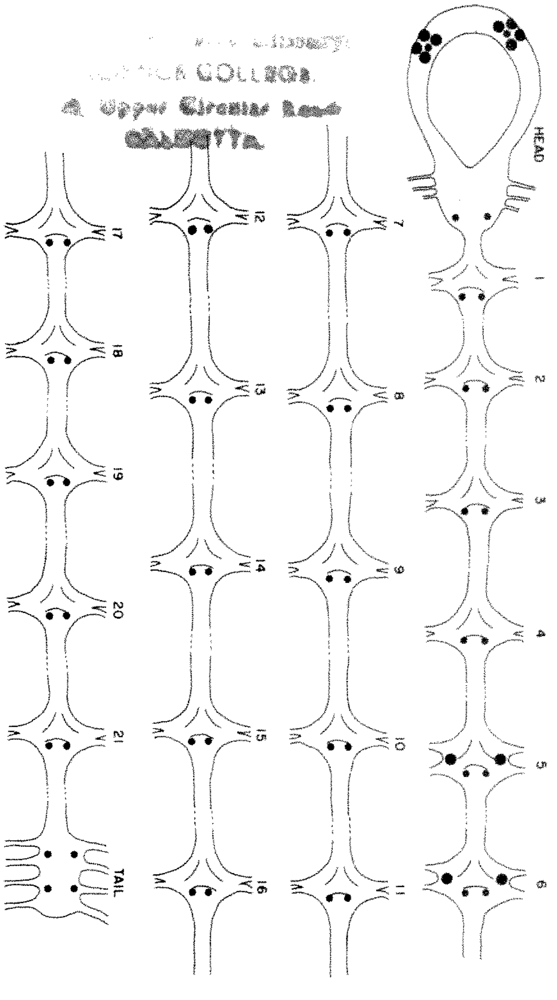


Fig. 3 This diagram illustrates the symmetry and repetitive organization of the nerve cord. The cell bodies stained by Lan3-1 in the 21 free midbody ganglia were verified in a large number of tests. The head and tail brains are more difficult to analyse and might contain additional neurones besides those indicated.

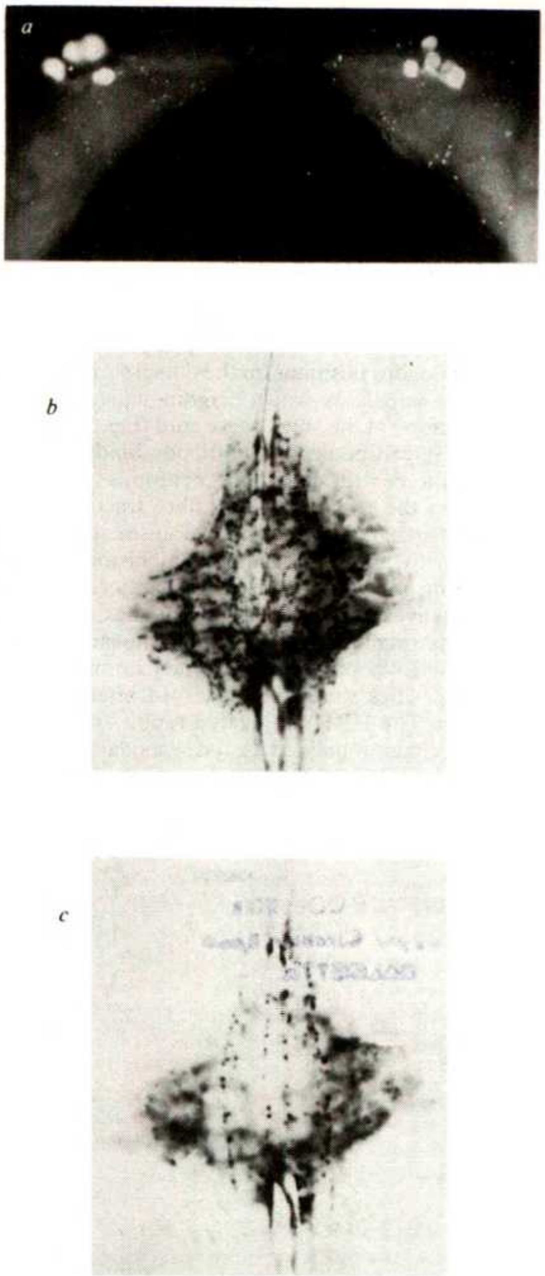


Fig. 4 Specific staining with Lan3-9. *a* Shows that this antibody also recognizes a bilateral set of five cell bodies in the supraoesophageal ganglion using rhodamine-conjugated second antibody. These differ from those stained by Lan3-1 (Fig. 2*a*). *b* And *c* show two focal planes through the same whole midbody ganglion stained by the PAP procedure (see Table 1 legend). No cell bodies can be seen here but several axons marked by discontinuous staining and a complex dense array of varicosities are visible.

also bound to Lan3-1, showing that this monoclonal antibody is an IgM.

All monoclonal antibodies were retested with fluorescently labelled secondary antibodies to confirm the staining pattern and detect features perhaps not visible with HRP. In the fluorescence procedure compromises have to be made between fixatives to achieve good labelling while maintaining low background. As the intact leech ganglion is 250-µm thick, an elevation in background fluoresence can critically impair the visibility of selectively stained features. Paraformaldehyde fixation has so far given the lowest background, making it possible to distinguish fine processes and varicosities at ×100 magnification. With Bouin's fixation, the background fluorescence tends to submerge fine processes buried in the neuropile while the visibility of superficial cell bodies remains excellent. However, with high power magnification the Lan3-1 varicosities can be seen with rhodamine-conjugated second antibody. In addition to Lan3-1, antibodies Lan3-3 and Lan3-4 recognize pairs of neurones in the midbody ganglion. Lan3-4 is specific for a pair of 40-µm cells situated between anterior and posterior roots, whereas Lan3-3 binds to a different pair closer to the anterior pole of the ganglion in the vicinity of the heart motor neurone.

Several antibodies that bind to homologous neurones in successive midbody ganglia bind to extra neurones in specialized ganglia, in the fused head ganglion or in the supraoesophageal ganglion. Ganglia 5 and 6 (G5, G6), containing almost twice as many neurones as the standard midbody ganglion, are specialized to subserve reproductive function and many of their supernumerary neurones innervate the genitalia. Lan3-1 recognizes pairs of large cells in G5 and G6 (Fig. 2*c*) in addition to the characteristic pairs of small cells (Fig. 2*b*). Previous studies[11] have shown that there is a pair of penile evertor motor neurones (PE) in G6 with the same position and axonal pattern as those stained by Lan3-1 in G6 and a homologous pair in G5 of unknown function. To demonstrate unequivocally that the large cells stained by Lan3-1 in G6 were the previously identified PE neurones, a double-labelling experiment using an injected dye was carried out. A cell body believed to be the left lateral PE cell was impaled with a microelectrode and its identity confirmed by its characteristic action potential (20 mV, non-overshooting) and its synaptic relationship with another major penile evertor cell, the rostral PE motor neurone[12]. Following electrophysiological identification, the fluorescent dye lucifer yellow[13] was injected through the recording electrode (Fig. 2*d*). To control for the effect of lucifer yellow on antibody staining, the left Retzius cell in the centre of the ganglion was also injected. After the ganglion had been fixed it was treated with hybridoma supernatant in the usual way and then treated with rhodamine-conjugated rabbit anti-mouse IgG. The results illustrated in Fig. 2*c* and *d* confirmed that the lateral PE cell contained the antigen to which the monoclonal antibody Lan3-1 binds. The same cell fluoresces yellow with the intrasomatically applied marker lucifer yellow (Fig. 2*d*) and red with the rhodamine-labelled mouse antibody (Fig. 2*c*). Of course, the monoclonal antibody bound to both the left injected lateral PE and the right uninjected cell. Under fluorescein isothiocyanate (FITC) optics, suitable for studying lucifer yellow fluorescence, a certain amount of rhodamine breakthrough is noticed in the right lateral PE cell. The lack of lucifer yellow interference with the antibody visualization is demonstrated by the absence of Retzius cell staining with rhodamine.

The subset of neurones labelled by Lan3-1 extends into the head and tail ganglion as well as into the supraoesophageal ganglion. The head and tail ganglia contain small pairs of neurones probably analogous to the small pair of cells that regularly repeat in the midbody ganglia. In addition, several large cells occur in the supraoesophageal ganglion. It is interesting that specific neurones in the supraoesophageal ganglion (Fig. 2*a*) with their separate embryological origin share an antigen with neurones in the rest of the nerve cord.

Of the 20 antibodies studied in detail, 2 bind to supernumerary neurones in the sex ganglia, G5 and G6. Like Lan3-1, the antibody Lan3-3 binds to extra neurones in G5 and G6 in addition to staining a pair of neurones which repeat in the midbody ganglia. The supernumerary cells labelled by Lan3-3 in the sex ganglia also lie between the anterior and posterior roots but on the dorsal surface whereas the lateral PE cell stained by Lan3-1 lies on the ventral surface. Axons and varicosities are not stained with Lan3-3.

Antibodies (for example, Lan3-12, 3-10, 3-9 and 3-11) were found that labelled cell bodies of neurones in the head, supraoesophageal ganglia or the anterior ganglia but not in any of the regularly repeating midbody ganglia. Lan3-12, 3-10 and 3-9 (Fig. 4*a*) all bind to neurones in the supraoesophageal ganglion. In addition, antigenically related processes occur throughout the nerve cord as fibres in the connective and nerve terminals in midbody ganglia are stained. As seen in Fig. 4*b* and *c* Lan3-9-

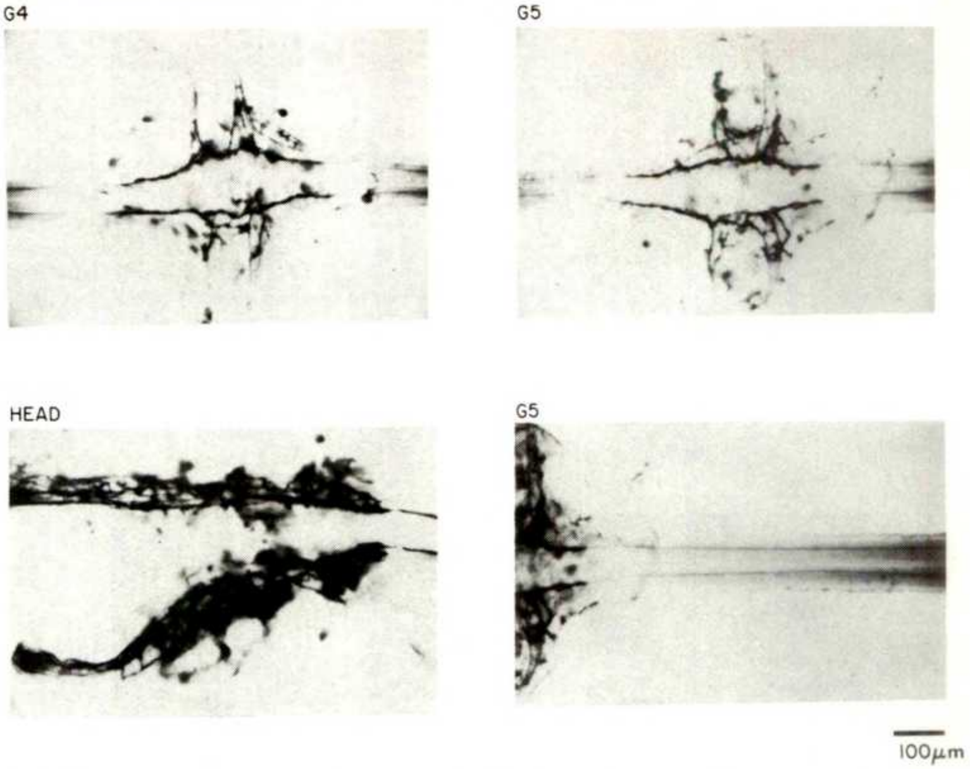


Fig. 5 The staining pattern of Lan2-3. Lan2-3 detects a single bilaterally symmetrical fibre system in each midbody ganglion. *a* and *b* show that major features of the fibre pattern repeat in adjacent ganglia. This uninterrupted staining of a branching fibre system contrasts with the punctate varicosities stained with both Lan3-1 (Fig. 2) and Lan3-9 (Fig. 4). The fibre system stained in the head ganglion is more complex (*c*). In *d* we show that a tight bundle of axons extends from a ganglion into the connective.

labelled axons run in the connective and similarly stained varicosities swamp the ganglionic neuropile.

Some antibodies stained only axons and neurites. Cell bodies have not yet been detected. An example is given in Fig. 5, where the staining pattern of Lan2-3 is seen to consist of relatively thick smooth fibres that pass through the connective (Fig. 5*d*) and ramify in the midbody ganglia (Fig. 5*a*, *b*) and head (Fig. 5*c*). This fibre system could have peripheral cell bodies, cell bodies in the tail ganglion or cell bodies devoid of antigen.

Nine monoclonal antibodies were isolated which stain more than one pair of neurones in a typical midbody ganglion. Six of these antibodies bind to mechanosensory neurones present in the typical 400-neurone midbody ganglion which selectively respond to touch, pressure and pain in clearly defined receptive fields on the skin of the leech. The typical mechanosensory cell body sends its axons (1) into the ipsilateral roots to innervate its own ganglion's receptive field and (2) via the ipsilateral connective into neighbouring ganglia to innervate partial receptive fields in adjacent body segments. These neurones have been extensively characterized electrophysiologically and through intracellular marker injections[1,6,7].

The four neurones responding to noxious mechanical stimuli (the nociceptive cells) all share a common antigen recognized by the monoclonal antibody Lan3-2 (Fig. 6*d*). This antigen is present in the cell body and all along the axonic and neuritic processes. It labels the ipsilateral projections and highlights a dense railroad-track-like pattern arising from the close packing of right and left axons running through the nerve cord. The nociceptive cells labelled by Lan3-2 and the pressure cells labelled by Lan3-5 (Fig. 6*c*) have been identified by double labelling (data not shown). The monoclonal antibody Lan3-5 labels the cell bodies and processes of three pairs of unidentified neurones and the two pairs of pressure cells. The pressure cells confine their axons to the ipsilateral ganglion and connective whereas the unidentified neurones cross to the contralateral side, allowing them the possible role of integrative interneurones. The antibody Lan3-6 binds to about 26 neurones in the typical posterior midbody ganglion including the pressure cells, possibly the touch cells and unidentified neurones (Fig. 6*b*). Additional antibodies that bind to pressure cells are Lan3-7 (Fig. 6*a*), Lan1-1 and Lan2-1 (not shown here). It is likely that these antibodies bind to different pressure cell antigenic determinants as they stain different sets of neurones.

Several antibodies show specific staining superimposed on a generalized uniform labelling. Examples are the brightly labelled roots of Lan3-5 (Fig. 6*c*) or the gradient of diaminobenzidine reaction product intensity of the Lan3-6-stained ganglion (Fig. 6*b*). Lan 3-6 reproducibly labels the constellation of neurones illustrated here, but in addition other cells, such as the Retzius cells, can be weakly labelled. Lan3-2 in Fig. 6*d* only binds to the four nociceptive cell bodies as visualized with the rhodamine-conjugated second antibody. If an HRP-conjugated second antibody is used, other cell bodies can show some degree of staining. Lan1-1 is an antibody that gives a somewhat higher degree of generalized staining in addition to binding to a definite constellation of ganglionic neurones. It also binds to many peripheral nerve bundles and cell bodies but not to muscle fibres or connective tissue. Specific cell staining is superimposed on the highest level of uniform cell body labelling with Lan3-8. This antibody is of special interest because it reproducibly stains the S-cell axon[1] as verified through double-labelling experiments. Occasionally, the S-cell body can be recognized but is generally submerged in the background. In contrast to Lan3-8, most of our antibodies that differentiate between specific nerve cells stain a small set of neurones. Of the 20 antibodies listed in Table 1, only 2 show strong general neuronal staining, 8 may give some general staining on which the reproducible specific staining is superimposed, but 10 show very high signal to noise labelling of only a few neurones. We have used a microphotometer to analyse the specificity of staining quantitatively. For example, in 1-µm thick sections of the supraoesophageal ganglion incubated with Lan3-9, the cell bodies similar to those shown in Fig. 4*a* had an optimal optical absorbance 50-fold greater than an adjoining unstained tissue. This makes us confident that we are detecting major differences in antigen concentration between neurones.

Discussion

We report here that monoclonal antibodies raised against the nervous system of the leech reveal an unexpectedly high level of antigenic diversity. Of the 475 hybridomas tested, approximately 300 bound to the leech nervous system. Of these, 41 different antibodies have been obtained which label specific subsets of neurones. Twenty of these antibodies have been analysed in detail and all give different staining patterns. These numbers suggest that many hybridomas secreting antibodies

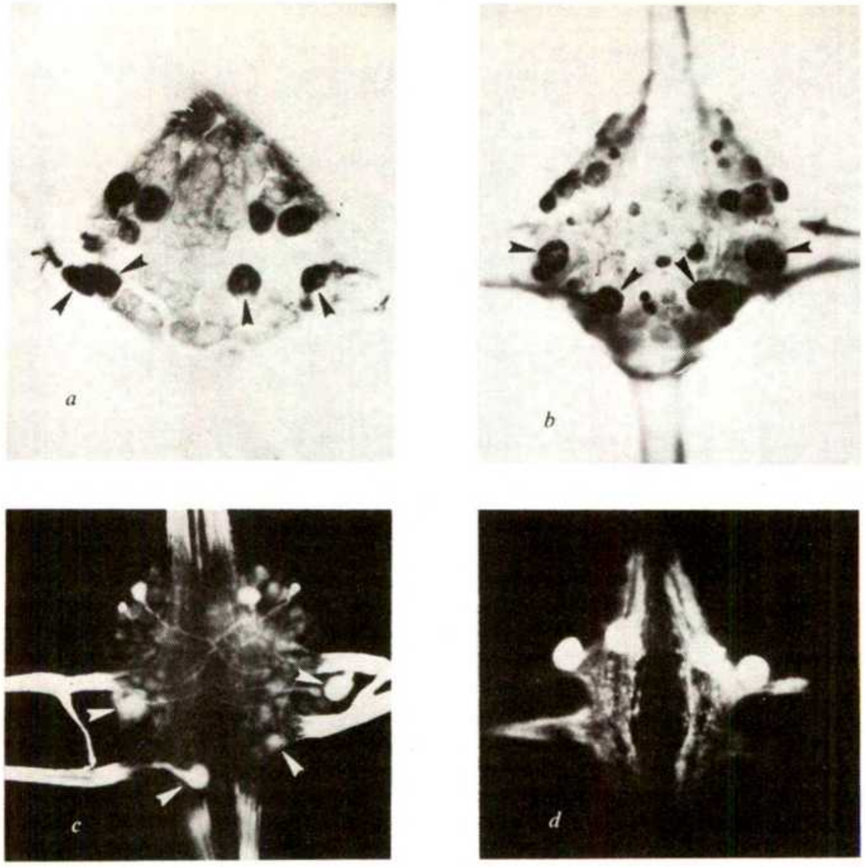


Fig. 6 Mechanosensory cells stained by four different antibodies. Three antibodies, Lan3-7 (*a*), Lan3-6(*b*) and Lan3-5 (*c*) stain the pressure cells (marked by arrows) as part of three distinct staining patterns. The pressure cells were identified by their morphology and in the case of Lan3-5 by double labelling with rhodamine-conjugated second antibody and lucifer yellow microinjection. In *d* the antibody Lan3-2 is shown staining the nociceptive cell bodies and processes.

against specific antigenic determinants have not yet been detected. The presence of large numbers of antigens restricted to small groups of neurones has implications for many aspects of neurobiology. Recent work on the rat retina has shown that monoclonal antibodies can distinguish between photoreceptors and glial cells[14]. The promise of this approach lies in the potential to identify molecules present in cells with a known electrophysiological function and neuroanatomical location. The chemical specificity of the antibodies then offers a bridge between electrophysiology, anatomy, the molecular nature and developmental behaviour of the antigen.

The potential of specific monoclonal antibodies as a tool in neurobiology can be seen, for example, in the antibody Lan3-1 that binds to a small group of cells in the leech nerve cord. The two large cells in the fifth and sixth midbody ganglia are known to be involved in the control of sexual behaviour. The cell in ganglion 6 was previously identified as one of the two major penile evertors, the lateral PE motor neurone[11,12]. The antibody staining technique has now confirmed the earlier impression that this type of neurone repeats in ganglia 5 and 6 but not in any other midbody ganglion. The obvious question raised by the staining pattern is the relationship between the large-diameter supernumary neurones in G5, G6 and the headbrains to the small-diameter neurones repeating in successive ganglia. Lan3-1 clearly does not code for general penile evertor function because it does not label the other major penile evertor cell, the rostral PE motorneurone, but it could still signify a sub-network controlling sexual function. Antibodies like Lan3-1 will serve as markers for studying leech development. By staining the nervous system at different developmental stages with Lan3-1 we may identify a particular sequence in which the cells express the antigen recognized by Lan3-1. In the case of Lan3-1, the special interest lies in relating these data to the sexual and hormonal development of the organism.

Considering the past work on the leech[6], it is of considerable importance that several antibodies were isolated that label the mechanosensory cells mediating the motor responses to the environmental stimuli touch, pressure and noxious mechanical stimulation. Such monoclonal antibodies will be helpful in extending the analysis of sensory systems in the leech where the detailed description of the first-order sensory neurones stands in contrast to our ignorance of higher-order sensory function. For example, Lan3-5 binds to all four pressure cells and to three other bilateral pairs of cells of so far unidentified function in the typical midbody ganglion. Conceivably, such immunocytochemical staining patterns could represent a network processing pressure sensation in which the nonidentified cells might be higher-order sensory neurones.

An integrated approach using many neurobiological techniques will be required to understand these staining patterns in detail. We have no experimental information on the molecular nature of the antigens showing restricted cellular distributions. It is clearly important to define molecules found in only a few neurones as we suppose that directly or indirectly they will lead us to the molecular mechanisms governing neuronal specificity.

We thank David Zipser for important contributions to this project and manuscript, Mike Ockler for drawing Fig. 1, David Lane for his help during our first fusion, Walter Stewart for the lucifer yellow dye, and James D. Watson for his encouragement. This work was supported by NSF grant BNS 78-248-72 and a Whitehall grant. During part of this project R. McK. was supported by an MRC fellowship.

Received 18 September; accepted 17 December 1980.

1. Muller, K. J. *Biol. Rev.* **54,** 99–134 (1979).
2. Hokfelt, T., Johansson, O., Ljungdahl, A., Lundberg, J. M. & Schutzberg, M. *Nature* **284,** 515–521. (1980).
3. Kohler, G. & Milstein, C. *Nature* **256,** 495–497 (1975).
4. Weisblat, D. A., Harper, G., Stent, G. S. & Sawyer, R. T. *Devl Biol.* **76,** 58–78 (1980).
5. Hagadorn, I. R. *Am. Zool.* **6,** 251–261 (1966).
6. Nicholls, J. G. & Baylor, D. A. *J. Neurophysiol.* **31,** 740–756 (1968).
7. Nicholls, J. G. & Purves, D. *J. Physiol., Lond.* **209,** 647–667 (1970).
8. Stuart, A. E. *J. Physiol., Lond.* **209,** 627–646 (1970).
9. Wilkinson, J. U. & Coggeshall, R. E. *J. comp. Neurol.* **162,** 397–416 (1975).
10. Zipser, B. *Nature* **283,** 857–858 (1980).
11. Zipser, B. *J. Neurophysiol.* **42,** 455–464 (1979).
12. Zipser, B. *J. Neurophysiol.* **42,** 465–475 (1979).
13. Stewart, W. W. *Cell* **14,** 741–759 (1978).
14. Barnstable, C. J. *Nature* **286,** 231–235 (1980).
15. Shulman, M., Wilde, C. D. & Kohler, G. *Nature* **276,** 269–270 (1978).
16. Sternberger, L. A. *Immunocytochemistry* (Prentice Hall, New Jersey, 1978).

Cloning of cDNA of major antigen of foot and mouth disease virus and expression in *E. coli*

Hans Küpper*, Walter Keller, Christina Kurz, Sonja Forss & Heinz Schaller

Department of Microbiology, University of Heidelberg, Im Neuenheimer Feld 230, 6900 Heidelberg, FRG

Reinhard Franze & Karl Strohmaier

Federal Research Institute for Animal Virus Diseases, Paul-Ehrlich-Str. 28, 74 Tübingen, FRG

Otfried Marquardt, Vladimir G. Zaslavsky & Peter Hans Hofschneider

Max Planck Institute for Biochemistry, Am Klopferspitz, Martinsried bei Munich, FRG

Double-stranded DNA copies of the single-stranded genomic RNA of foot and mouth disease virus have been cloned into the Escherichia coli *plasmid pBR322. A restriction map of the viral genome was established and aligned with the biochemical map of foot and mouth disease virus. The coding sequence for structural protein VP1, the major antigen of the virus, was identified and inserted into a plasmid vector where the expression of this sequence is under control of the phage λ P$_L$ promoter. In an appropriate host the synthesis of antigenic polypeptide can be demonstrated by radioimmunoassay.*

FOOT AND MOUTH disease virus (FMDV) is a picornavirus which affects cloven-footed animals. Like other picornaviruses its genome consists of a single-stranded RNA molecule of about 8,000 bases. It is polyadenylated at its 3′ end, has a poly(C) tract close to the 5′ end and a protein covalently attached to its 5′ end. Although it does not have the usual eukaryotic mRNA cap it serves as a mRNA both *in vivo* and *in vitro*. Only the region between the poly(C) tract and the 3′ end of the RNA carries the coding sequence, with a single initiation point for translation close to the poly(C) tract (for reviews see refs 1, 2). FMDV virions contain primarily four capsid proteins, identified as VP1, VP2, VP3 and VP4 (terminology as in ref. 1 is used, VP1 corresponds to VP3 in ref. 2) which are synthesized as a common polypeptide precursor in the infected cell and map close to the 5′ end of the viral genome[3]. Of these proteins VP1 is of particular interest because it carries epitope(s) inducing neutralizing antibodies[4,5]. For this reason it probably contributes to the high antigenic variation of FMDV which results in 7 serotypes and more than 60 subtypes. This variability causes major problems in vaccination programmes and makes the virus interesting not only for academic but also for applied research.

Because of recent advances in DNA technology a double-stranded DNA copy of the viral RNA would be of advantage for further detailed analysis of the FMDV genome, and, in particular, of the regions coding for the structural proteins. Thus, the antigenic variability could be analysed on a molecular basis by restriction mapping and DNA sequencing. The synthesis of VP1 in a bacterial host would be of obvious interest; although this protein is a less potent immunogen than the intact virus, it can protect animals from infection, as has been shown for swine[6]. Here we report the isolation of cDNA clones covering 70% of the FMDV genome and present a restriction map for the genome of FMDV strain O$_1$K and the precise location of capsid protein VP1 within this map. We also show that the VP1 coding sequence is expressed in *E. coli*.

Physical and genetic map of FMDV, serotype O$_1$ Kaufbeuren

Single-stranded cDNA was synthesized from FMDV RNA using oligo(dT)$_{12-18}$ as a primer for AMV reverse transcriptase. This cDNA was rendered double stranded by *E. coli* DNA polymerase I in the presence of the four [α-^{32}P]-labelled deoxynucleoside triphosphates. The size of the double-stranded cDNA was determined by electrophoresis on a 1% agarose gel using appropriate markers. More than 40% of the cDNA synthesized in this way was between 4,500 and 8,000 base pairs long. As shown in Fig. 1 a major fraction of the cDNA was much

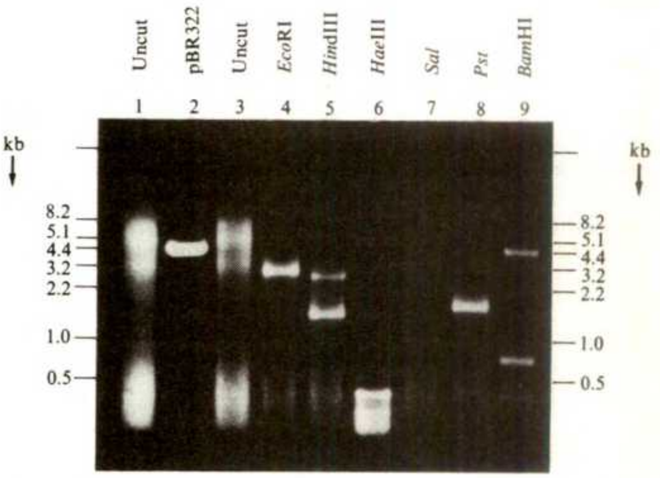


Fig. 1 Restriction analysis of FMDV cDNA. Strain O$_1$K was isolated in Kaufbeuren during an epizootic outbreak that occurred in southern Germany in 1966–67. The virus was cloned three times from single plaques and passaged about 50 times in cell cultures. The identity of the passaged virus was confirmed by complement binding assays with the original antibodies. For RNA isolation the virus was grown in BHK cell monolayers in several roller bottles. Viral RNA was prepared essentially as described by Grubman *et al.*[17]. Single-stranded cDNA was prepared by AMV reverse transcriptase (80 units) initiated by 0.1 unit (A_{260}) of pT$_{12-18}$ primer in 100 µl of 50 mM Tris-HCl *p*H 8.3, 50 mM KCl, 8 mM $MgCl_2$, 25 mM β-mercaptoethanol, dATP, dCTP, dGTP and dTTP each at 0.5 mM and 14 µg FMDV RNA having been pretreated with 2.5 mM CH_3 HgOH for 3 min at room temperature. After incubation at room temperature for 5 min and at 42 °C for 60 min, the mixture was boiled in a water bath for 3 min and quickly cooled to dissociate the cDNA–RNA hybrids. The cDNA strand was rendered double stranded by combining the above mixture with 100 µl of 0.2 M HEPES buffer *p*H 6.9, 90 mM KCl, 0.4 mM each of dATP, dCTP, dGTP and dTTP, 50 µCi each of the four ^{32}P-labelled deoxytriphosphates and 50 units DNA polymerase I, before incubation at 15 °C for 2 h. To stop the reaction, the mixture was extracted with phenol and passed through a Sephadex G-150 column. The double-stranded cDNA was further analysed by digestion with various restriction enzymes. cDNA (2×10^4 c.p.m.) was mixed with 7 µg of adenovirus 2 (Ad2) DNA and digested with the enzymes indicated in lanes 4–9. Fragments were analysed on a 1.0% agarose gel. Staining with ethidium bromide showed the position of the Ad2 DNA fragments which were used as internal markers and were also an indication for complete digestion. cDNA fragments were detected by autoradiography. Linearized pBR322 was added as a marker (lane 2). kb, Kilobase.

---

* Present address: Biogen SA, 3 rte de Troinex, 1227 Carouge, Geneva, Switzerland.

0028–0836/81/070555–05$01.00

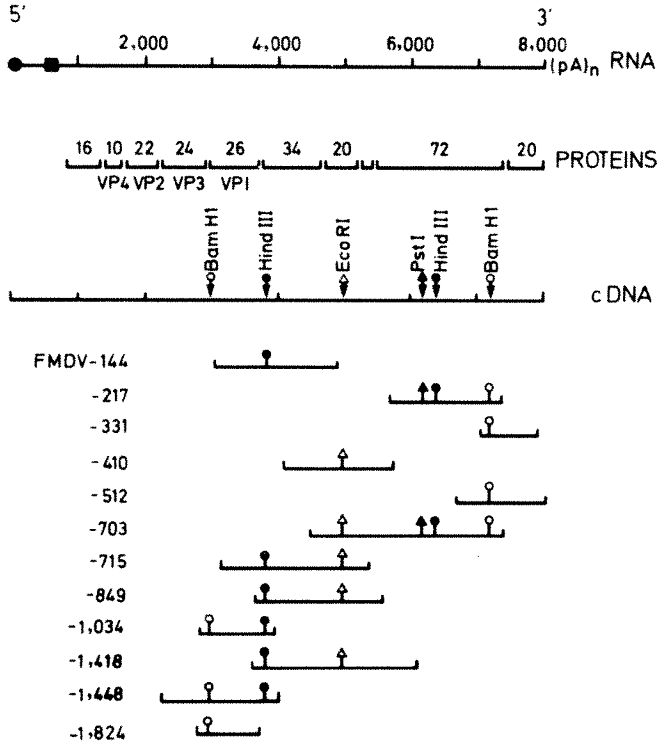


Fig. 2 Localization of cloned FMDV cDNA fragments. For convenience the size of the RNA is given as 8,000 bases instead of the 7,800 bases estimated by physical analysis[18] and the *Bam*HI restriction site is referred to as map position 3,000. The molecular weights ($\times 10^{-3}$) of the proteins are as indicated. VP1 to VP4 are the viral capsid proteins. Within the RNA molecule ■ indicates the internal poly(C) stretch close to the 5' end and ● indicates the viral genomic protein (VPg) covalently linked to the 5' end of the viral RNA. Plasmid pBR322 was digested completely with *Pst*I endonuclease. The mixture was extracted with 1 vol phenol and precipitated with 2.5 vol ethanol. Homopolymeric dG-tails were added using terminal deoxynucleotidyl transferase in a 50-µl reaction volume containing 1 µg of cleaved plasmid, 200 mM Na-K-cacodylate (pH 7.2), 1 mM $CoCl_2$, 1 mM β-mercaptoethanol, 200 µM dGTP and 12 units of enzyme. Incubation was for 30 min at 37 °C. Approximately 5 µg of double-stranded cDNA were elongated with dCTP residues by terminal deoxynucleotidyl transferase in a 50-µl reaction volume as above, but containing 100 µM dCTP and 15 units of enzyme. Incubation was for 5 min at 37 °C, and the reaction was stopped by freezing. For annealing, 20 ng dG-tailed *Pst*I-cleaved pBR322 vector and 50 ng dC-tailed double-stranded cDNA were mixed in 30 µl 1 × SSC buffer. The mixture was placed in a water bath at 69 °C and allowed to cool slowly overnight to 35 °C. The recombinant DNA mixture was then used to transform competent *E. coli* HB101 cells. The analysis of the plasmids from colonies showing Tet^r and Amp^s phenotype and the deduction of the cDNA restriction map from the experiment described in Fig. 1 is given in the text. The experiments were carried out in L3 B1 conditions in accordance with the German guidelines for recombinant DNA research.

shorter. This bimodal distribution may be explained by a high degree of secondary structure in certain parts of the RNA template. For further analysis the cDNA was digested with various restriction endonucleases (Fig. 1, lanes 4–9). Simple restriction patterns were obtained with restriction endonucleases *Eco*RI and *Pst*I or *Hin*dIII and *Bam*HI, which generated one or two restriction fragments, respectively.

A restriction map can be deduced using the size of these fragments and the relative intensity of these bands in a gel. 3'-Proximal cDNA sequences occur more often in the cDNA population and the corresponding bands are therefore more intense. Because the 5'-proximal cDNA sequences are heterogeneous in length no corresponding fragments can be detected. Of two fragments generated by *Hin*dIII the more intense smaller fragment must contain sequences closer to the 3' end of the FMDV genome. For the *Bam*HI digest the intensities of the two bands are about equal; however, because of the uniform labelling of the cDNA the 4-kilobase band comprises less molecules than the smaller (0.8 kilobase) band. In this way a restriction map of the FMDV genome was obtained and aligned relative to the 3' end of the RNA (Fig. 2). Approximate correlations between the different viral proteins and the viral genome were deduced from the order of translation of the different polypeptides and from their apparent molecular weights[7]. This correlation can now be used to predict the position of viral genes within specific restriction fragments of the FMDV genome. This prediction is, of course, approximate but as shown below it was successfully applied to find the coding region for VP1. The predicted position was, in fact, only 300 base pairs away from the position shown in Fig. 2 (between the *Bam*HI cut at map position 3,000 and the *Hin*dIII cut at map position 3,847).

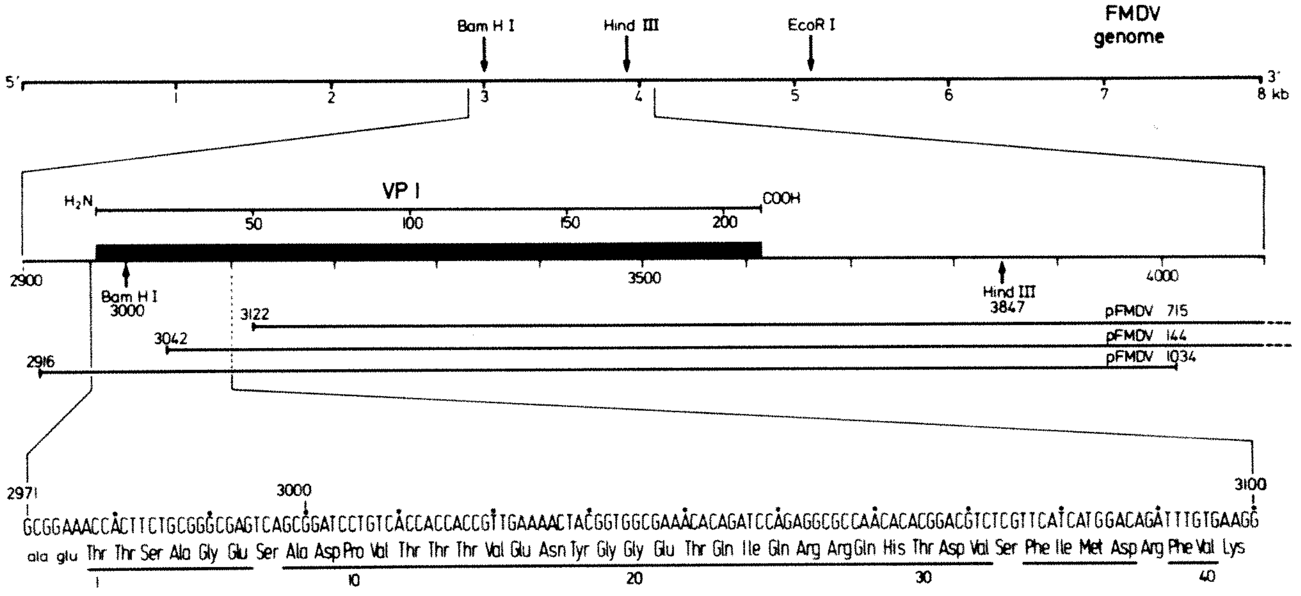

Bam H I
Hind III
EcoR I
FMDV genome
5'
3'
1
2
3
4
5
6
7
8 kb
VP I
H₂N
COOH
50
100
150
200
2900
3500
4000
Bam H I 3000
Hind III 3847
3122
3042
2916
pFMDV 715
pFMDV 144
pFMDV 1034
2971
3000
3100
GCGGAAACCACTTCTGCGGGCGAGTCAGCGGATCCTGTCACCACCACCGTTGAAAACTACGGTGGCGAAACACAGATCCAGAGGCGCCAACACACGGACGTCTCGTTCATCATGGACAGATTTGTGAAGG
ala glu Thr Thr Ser Ala Gly Glu Ser Ala Asp Pro Val Thr Thr Thr Val Glu Asn Tyr Gly Gly Glu Thr Gln Ile Gln Arg Arg Gln His Thr Asp Val Ser Phe Ile Met Asp Arg Phe Val Lys
1
10
20
30
40

Fig. 3 Exact correlation of the physical map of the FMDV genome with cloned cDNA sequences that code for VP1. The position of the VP1 gene, of cloned FMDV sequences and of unique restriction sites in the enlarged part of the genome are based on the nucleotide sequence of pFMDV-1034 (ref. 10). The nucleotide sequence at the 5' end of the VP1 gene and the deduced NH_2-terminal amino acid sequence for VP1 are presented in the second enlargement. Amino acids that agree with those determined experimentally[9] are underlined.

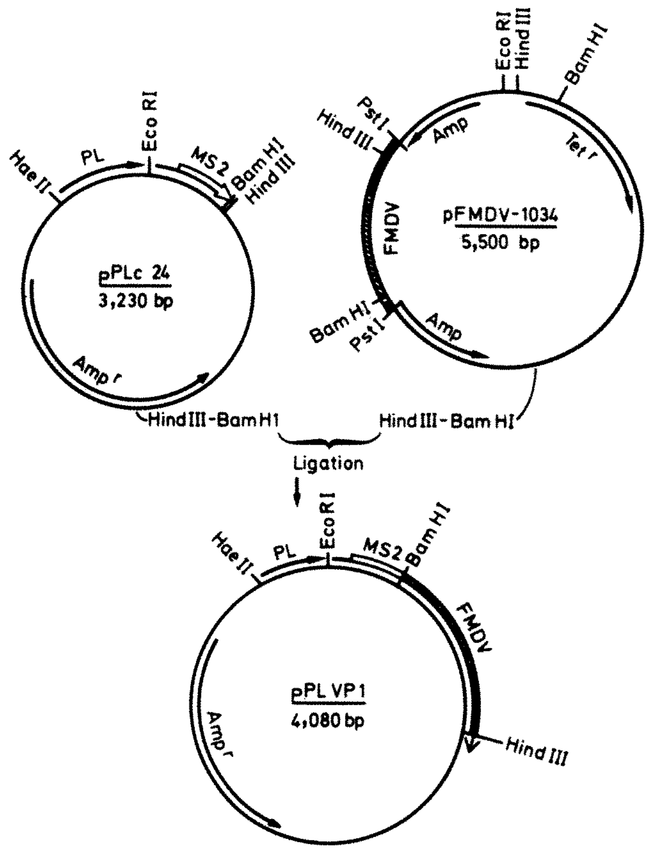


Fig. 4 Construction of plasmid pPL VP1. The schematic outline shows the vector plasmid pPL c24 and the donor plasmid pFMDV-1034. Amp, β-lactamase region coding for resistance to ampicillin; Tet, coding region for resistance to tetracycline; PL, λ P_L control element; MS2, DNA fragment derived from bacteriophage MS2 where the boxed part indicates the coding sequence for the first 99 amino acids of the MS2 polymerase; FMDV, insert derived from FMDV cDNA, the shadowed area shows the coding region for VP1. The predicted fusion polypeptide from the constructed plasmid pPL VP1 is indicated. Only restriction sites relevant for this construction are shown. The FMDV insert of pFMDV-1034 was isolated by electrophoresis on a 1% agarose gel after digestion of the plasmid with *Pst*I. The isolated FMDV fragment was digested with *Hin*dIII and *Bam*HI, the vector was sequentially digested with *Bam*HI and *Hin*dIII. 0.2 pmol of restricted vector were combined with 0.1 pmol of the restricted FMDV fragment in 50 µl ligation buffer (50 mM Tris-HCl *p*H 8.0, 10 mM Mg Cl_2, 50 mM NaCl, 0.2 mM EDTA, 1 mM ATP and 1 mM dithioerythritol) and incubated for 16 h at 15 °C after addition of T4 ligase. The most likely construction created in this way is pPL VP1. The *Bam*HI–*Hin*dIII fragment excised from pPL c24 is only 10 nucleotides long with 6 nucleotides overlapping and will not be efficiently reinserted. Enzymatic reactions were stopped by incubation at 65 °C for 10 min.

Cloning of FMDV cDNA into pBR322 and characterization of the hybrids

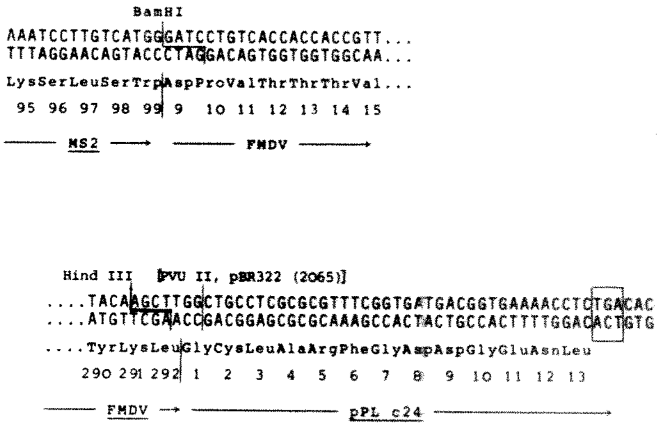
To insert the double-stranded cDNA molecules into plasmid pBR322, oligo(dC)-tails were added to their 3′ termini after treatment with nuclease S_1 (ref. 8). The dC-tailed DNA molecules were annealed to pBR322 DNA to which 3′ oligo(dG) sequences had been added after cleavage with *Pst*I. The annealed DNA was then used to transform competent cells of *E. coli* HB101. One thousand transformants which were ampicillin sensitive and tetracycline resistant were transferred to microtitre plates. Replicas of each plate were used for hybridization with ^{32}P-labelled fragmented viral RNA (average chain length 300 nucleotides). Colonies which displayed strong hybridization with the ^{32}P-labelled probe were selected for further analysis assuming that these contain large inserts.

Plasmids were isolated and the FMDV DNA inserts excised with *Pst*I. Surprisingly, none of the clones contained inserts longer than about 2,500 base pairs. Further analysis of the FMDV inserts by restriction mapping of the *Pst*I-cleaved DNA with *Eco*RI, *Bam*HI and *Hin*dIII showed that the cloned DNA did not all begin with overlapping sequences at the 3′ end as expected from the oligo(dT)-primed synthesis of the cDNA, but originated from various parts of the FMDV genome and covered about 75% of the 8,000 bases (Fig. 2). Predicted overlaps between various inserts were further confirmed by comparative structural analysis using restriction nucleases *Hae*III, *Hin*fI, *Hha*I, *Taq*I and *Ava*I (results not shown). These results also confirmed the physical map of the FMDV genome which was based originally on the use of only four restriction enzymes. In conjunction with the biochemical map of the viral proteins this predicts that several clones (for example FMDV-144 and FMDV-715) should contain nucleotide sequences which code for capsid protein VP1 (Fig. 2). The reasons for the unexpectedly small size of the FMDV cDNA inserts and their scattering throughout the genome are not understood. Before cloning, 40% of the cDNA molecules were >4,000 base pairs long starting from the 3′ end of the FMDV genome, as demonstrated by the sharpness of the digestion patterns with various restriction endonucleases (Fig. 1). Nevertheless, the average size of the cloned cDNA inserts was only about 1 kilobase, even when only cDNA molecules larger than 4 kilobases were used (isolated by preparative polyacrylamide gel electrophoresis, data not shown). This rules out the possibility that the result obtained was due to nonspecific degradation of the cDNA during its synthesis and suggests that oligo(dC)-tails were added internally by the terminal transferase at single-stranded nicks, which may have been introduced during the preparation of the cDNA.

In all clones analysed both terminal *Pst*I sites were conserved and analysis of overlapping cloned FMDV DNA showed complete agreement of the nucleotide sequence within the overlapping DNA part. This again suggests that the short size of the inserted FMDV DNA is not due to internal recombination or deletion.

Identification of DNA inserts coding for VP1

By comparison with the original biochemical map, clones FMDV-715 and FMDV-144 seemed to contain the coding sequence for viral protein VP1. Nucleotide sequences from the corresponding parts of the cloned cDNA were therefore determined and compared with the known sequence of the NH_2-terminal 40 amino acids of VP1[9]. A fit was obtained with the 5′-proximal end of the DNA insert of pFMDV-144, which contained a sequence capable of coding for amino acids 22–40 of protein VP1 (Fig. 3). This indicated that the VP1 gene maps about 300 nucleotides to the left of the position estimated from our original map. Additional clones were therefore selected which also contained the missing part of the VP1 gene (see below). Of these, the insert from pFMDV-1034 was analysed in detail and its nucleotide sequence determined[10]. It showed that the insert from pFMDV-1034 extends 126 nucleotides into the 5′ direction beyond the end of clone pFMDV-144. After the first



```
               BamHI
AAATCCTTGTCATGGGATCCTGTCACCACCACCGTT...
TTTAGGAACAGTACCCTAGGACAGTGGTGGTGGCAA...
LysSerLeuSerTrpAspProValThrThrThrVal...
 95 96 97 98 99 9  10 11 12 13 14 15
—— MS2 ——→  ———— FMDV ————→
```

```
    Hind III  [PVU II, pBR322 (2065)]
....TACAAGCTTGGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACAC
....ATGTTCGAACCGACGGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTG
....TyrLysLeuGlyCysLeuAlaArgPheGlyAspAspGlyGluAsnLeu
    290 291 292 1  2  3  4  5  6  7  8  9  10 11 12 13
—— FMDV → ———————— pPL c24 ————————→
```

Fig. 5 Junction sequences at the borders of the inserted FMDV DNA in plasmid pPl VP1, as predicted by the construction outlined in Fig. 4. The nucleotide sequence at the *Bam*HI site was confirmed by DNA sequencing. The corresponding amino acid sequence was deduced from the DNA sequence.

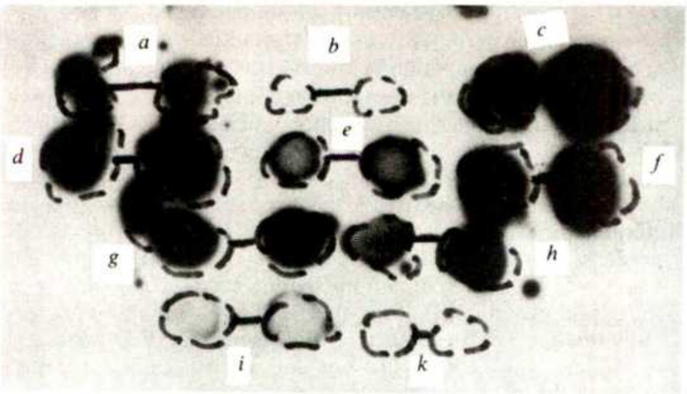


Fig. 6 Autoradiograph showing the detection of VP1-specific material in bacterial extracts. Colonies of *E. coli* K12ΔH1 (K12 M72 Lac_{am} ΔtrpEA2 strep^{r}[λC1 857 N_{am7} N_{am53} ΔH1 bio^{-}])[14] transformed by pPl VP1-1 to pPl VP1-7 or by pBR322 and pPl c24 as controls were grown at 27 °C and served as inoculum for liquid cultures (100 ml, L-broth). Cultures were grown up to 6×10^8 cells ml^{-1} at 27 °C and then the temperature was shifted and kept for 5 h at 42 °C. After chloroform treatment the bacteria were collected by centrifugation, resuspended in 1 ml phosphate-buffered saline (PBS) and sonicated for 1 min in an ice bath. The bacterial debris were washed with 0.5 ml PBS and both supernatants were combined. 5 μl of these extracts were applied in duplicate onto polyvinyl disks coated with anti-O_1K IgG prepared from hyperimmune serum of guinea pigs. The disks were incubated for 4 h at 37 °C and washed as described elsewhere[15] except that 0.05% NP40 was added to the washing solution. Finally, the disks were incubated overnight at 4 °C with ^{125}I-labelled anti-VP1 (4×10^6 c.p.m. per μg; 5×10^6 c.p.m. per disk, preabsorbed with 0.5 ml extract from K12 ΔHI/pPL c24). After thorough washing the disks were autoradiographed. *a* and *d–i*, extract prepared from *E. coli* K12ΔH1 carrying plasmid pPl VP1-1 to pPL VP1-7, respectively. *b*, *k*, Extract from K12ΔH1/pPL c24 and K12ΔH1/pBR322 as negative controls. *c*, FMDV virus (2 μg) as positive control.

61 nucleotides the sequence is in agreement with the N′-terminal amino acid sequence of VP1 (the sequence relevant to this discussion is shown in Fig. 3). This sequence is followed up to the end of pFMDV-1034 (map position 4,010) by a continuous translation reading frame and thus codes for more than 350 amino acids of FMDV-specific polypeptide which is well above the estimated size of VP1[1,2,11]. Except for the difference of one nucleotide, identical sequences were obtained from corresponding segments of the independently isolated clones FMDV-144 and FMDV-715, which start at map positions 3,041 and 3,122, respectively (Fig. 3). Both the identity of the common nucleotide sequence between independent isolates and the presence of a continuous open reading frame indicate that this part of the FMDV genome was cloned intact and that it contains the whole coding sequence for VP1. This conclusion is further supported by the results obtained for expression of this sequence in *E. coli* (see below).

Specific screening for genes coding for viral proteins

Having localized the gene for VP1 within a specific part of the FMDV genome, the whole library can now be screened specifically for DNA inserts carrying the coding sequence for viral proteins. The *Pst*I–*Hin*dIII fragment of pFMDV-144, which carries most of the coding sequence for VP1, was used as a probe to select other clones having inserts covering the same part of the FMDV genome. From the total library of 1,000 clones, 16 displayed strong hybridization and were analysed further by *Bam*HI restriction mapping and size determination of the resulting fragments. The presence of a *Bam*HI restriction site within the insert and the size of the insert make it possible to identify clones which carry the coding sequence not only for VP1 but also for all or a portion of the nucleotide sequence of VP2 and VP3. Clones selected in this way were identified as FMDV-1034, FMDV-1448 and FMDV-1824 (Fig. 2). In an analogous screening process the *Bam*HI–*Hin*dIII fragment of pFMDV-217 was used to probe for sequences originating from the 3′ end of the genome. Analysis of the hybridization data confirmed that the localization of the clones shown in Fig. 2 was correct. No clone hybridized to both parts of the genome, which again indicates that there were no recombinant DNA molecules carrying the entire FMDV genome.

Construction of plasmids expressing VP1 in *E. coli* cells

To express the viral protein VP1 in a bacterial host we chose an expression vector carrying the strong leftward promoter of bacteriophage λ the activity of which can be controlled by using host strains that carry the gene for a temperature-sensitive λ repressor (*c*I *ts*). This vector (pPL c24, Fig. 4, ref. 12) has been used successfully to express the gene for human fibroblast interferon[13]. It carries a unique restriction site for *Bam*HI at amino acid 99 of the MS2 replicase gene and a unique site for *Hin*dIII 10 nucleotides downstream. Using these two sites, the *Bam*HI–*Hin*dIII fragment of pFMDV-1034 can be inserted in the vector such that only one orientation of insertion is possible and the information for the MS2 gene will be in phase with the coding sequence of FMDV starting at amino acid 9 of coat protein VP1 (Figs 3, 4). Thus, the MS2 sequences can be used to promote polypeptide synthesis of the coat protein VP1, which as a processing product does not have its own initiation signals. A fusion polypeptide of 396 amino acids should be synthesized carrying the N-terminal 99 amino acids of MS2 replicase, 284 amino acids of FMDV and 13 amino acids coded for by the vector. The junction sequences between the MS2 replicase gene and the FMDV DNA at the *Bam*HI site and between the FMDV DNA and the vector DNA at the *Hin*dIII site are shown in Fig. 5.

Hybrid plasmids were constructed as outlined in Fig. 4 and transformed into an *E. coli* C600 r_{k}^{-} m_{k}^{-} (λ) host. Plasmids were isolated from seven colonies which showed a positive response by hybridization to a labelled FMDV probe and analysed for expression. For this purpose the plasmids were retransformed into an *E. coli* K12 ΔHI host[14], where expression from the P_{L} promoter can be controlled by temperature. Plasmids pPL c24 and pBR322 were included as controls.

From each transformation isolates were assayed for VP1-specific antigenic determinants in bacterial extracts (Fig. 6), essentially as described by Broome and Gilbert[15]. Positive results were obtained only with transformants carrying the hybrid plasmids but not with those carrying the vector plasmid pPL c24 (Fig. 6*b*) or pBR322 (Fig. 6*k*). Varying intensities of the positive signals may be due to variation in the preparation of extracts as the inserted plasmids should be the same in the different isolates.

The specificity of the antigenic material in the bacterial extracts was further demonstrated in a series of additional experiments which will be reported in detail elsewhere. An indirect enzyme-linked immunosorbent inhibition assay[16] using mouse anti-VP1 serum confirmed the results reported above. In the radioimmunoassay different sets of experiments showed that the antigenic material produced in the bacterial cell competes efficiently with ^{125}I-labelled FMDV, and that it did not react with labelled mouse anti-VP1 antibodies if these were exposed to FMDV before the test. A 60-fold dilution of the bacterial extract used in Fig. 6*d* was still positive when tested for VP1 in the radioimmunoassay. This indicates that at least 1,000 molecules of antigenic polypeptide are synthesized per cell. Extracts prepared from an *E. coli* C600 r_{k}^{-} m_{k}^{-}(λ) host carrying the plasmid pPL VP1 gave a negative result in the radioimmunoassay, indicating that the synthesis of antigenic material is under control of the P_{L} promoter. Finally, expression experiments in minicells showed that the protein displaying specific antigenicity has the size of the predicted fusion polypeptide.

Taken together, these data suggest that viral protein VP1 can be synthesized in an antigenically active state in *E. coli* cells, and thus may provide a new way of producing a vaccine without the risk of accidental infection by incompletely inactivated virus. The genes for VP1 of other serotypes are of equal interest and should be made available by the same techniques used in the present work. Comparison of the nucleotide sequences specific to the VP1 coding regions might give some insight into the molecular mechanism of antigenic variability in FMDV.

We thank Dr M. Mussgay for his interest and support, E. Bürgelt and M. Kauzmann for technical assistance, and Dr E. Domingo for helpful discussion.

Received 6 October; accepted 2 December 1980.

1. Sangar, D. V. *J. gen. Virol.* **45,** 1–13 (1979).
2. Bachrach, H. L. in *Beltsville Symp. Agricultural Research* (ed. Romberger J. A.) 3–32 (Allanheld Osmun, Montclair, 1977).
3. Sangar, D., Black, D. N., Rowlands, D. J. & Brown, F. *J. gen. Virol.* **35,** 281–297 (1977).
4. Laporte, J. *et al. C. r. hebd Séanc. Acad. Sci., Paris* **D276,** 3399–3402 (1973).
5. Bachrach, H. L., Moore, D. M., McKerchner, P. D. & Polatnick, J. *J. Immun.* **115,** 1635–1641 (1975).
6. Bachrach, H. L., Moore, D. M., McKerchner, P. D. & Polatnick, J. *Perspectives Virol.* **10,** 147–159 (1978).
7. Doel, T. R., Sangar, D. V., Rowlands, D. J. & Brown, F. *J. gen. Virol.* **41,** 395–404 (1978).
8. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3727–3731 (1978).
9. Strohmaier, K., Wittmann-Liebold, B. & Geissler, A. -W. *Biochem. biophys. Res. Commun.* **85,** 1640–1645 (1978).
10. Kurz, C., Forss, S., Küpper, H., Strohmaier, K. & Schaller, H. *Nucleic Acids Res.* (submitted).
11. Vande Woude, G. F. & Bachrach, H. L. *J. Virol.* **7,** 250–259 (1971).
12. Remaut, E., Stanssens, P. & Fiers, W. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
13. Derynck, R. *et al. Nature* **287,** 193–197 (1980).
14. Bernard, U. *et al. Gene* **5,** 59–76 (1979).
15. Broome, S. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2746–2749 (1978).
16. Abu Elzein, E. M. E. & Crowther, J. R. *J. Hyg., Camb.* **80,** 391–399 (1978).
17. Grubman, M. J., Baxt, B. & Bachrach, H. L. *Virology* **97,** 22–31 (1979).
18. Danoya, D. C. *et al. Virology* **84,** 230–235 (1978).

Generation of F_1 hybrid cytotoxic T lymphocytes specific for self H–2

Keiichiro Nakano, Ichiro Nakamura & Gustavo Cudkowicz

Departments of Pathology and Microbiology, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Cytotoxic T lymphocytes specific for self H–2 antigens are generated by murine F_1 hybrid (H–2 heterozygous) spleen cells cultured with irradiated parental (H–2 homozygous) splenocytes. The effectors bind to heterozygous and homozygous cells bearing the appropriate H–2 alleles but only lyse homozygous targets. Autoreactivity for membrane-bound molecules of normal cells may be a mechanism for regulating cellular interactions.

THYMUS-DERIVED lymphocytes are able to recognize autologous membrane-bound molecules coded for by the major histocompatibility complex (MHC). In the mouse, such recognition of self is the basis for the restriction of cell-mediated immune responses against viruses[1], haptens[2–4] and minor histocompatibility antigens[5] by H-2K, H-2D, H-2L and H-2I molecules. Moreover, it is through the recognition of self H-2I molecules that macrophage–lymphocyte interactions lead to immune responses by T and B lymphocytes[6,7]. In spite of the pervasiveness of self MHC recognition in immunology (see a recent discussion by Dausset and Contú[8]), mature effector cells or immunoglobulins specific for, rather than restricted by, self MHC molecules are generally not demonstrable. Regulatory mechanisms presumably prevent or control precursor lymphocyte differentiation and maturation into effectors that would entail autoimmune disorders. However, autoreactivity may be a double-edged sword: if directed against relevant cell-surface molecules such as the MHC products, anti-self effectors may serve to modulate cellular activities and maintain, instead of disrupt, homeostasis. During *in vitro* studies on the induction of F_1 hybrid anti-parent cell-mediated immunity, we have been able to demonstrate the existence of cytotoxic T lymphocytes (CTLs) specifically sensitized against self H–2 molecules. These T lymphocytes were unique in that they were capable of binding to target cells carrying the appropriate H–2 genes in single (heterozygous) or double (homozygous) dose. Most interestingly, the autoreactive CTLs were capable of lysing the homozygous but not the heterozygous targets.

Induction of F_1 anti-parent CTLs

F_1 anti-parent cytotoxic lymphocytes were generated in (C57BL/6×DBA/2)F_1 anti-C57BL/6 (abbreviated B6D2F_1 anti-B6) mixed spleen-cell cultures and tested for cytolytic activity by the release of radioactive ^{51}Cr from prelabelled target cells[9–12]. F_1 hybrid lymphocytes bearing the Thy-1 surface antigen[11], and thus of the T-lymphocyte class, became strongly cytotoxic for thioglycollate-induced peritoneal exudate cells (PECs) of the stimulating parental strain but not for syngeneic F_1 target PECs or for cells of the second parental strain (Fig. 1*a*). The susceptibility of F_1, DBA/2 and B6 targets to immune cytolysis was compared by raising allogeneic CTLs of appropriate specificity and measuring cytolysis simultaneously (Fig. 1*b*, *c*). As all the target cells were specifically lysed by appropriate allogeneic CTLs, the exclusive lysis of B6 PECs by F_1 anti-parent CTLs probably reflects their immunological specificity. Target determinants detected on PECs were also expressed on parental lymphoblasts stimulated by mitogens and on parental lymphoma cells[10,12].

F_1 spleen cells cultured for 5 days without parental stimulators failed to generate specific anti-parent CTLs even though the culture medium contained prescreened fetal calf serum. Low levels of cytotoxic activity were sometimes detectable among such F_1 cells cultured in the presence of fetal calf serum (for example, 10–15% specific lysis), but in these cases cytolysis exhibited no preference for PECs of either parental type and was polyclonal in nature. Thus the induction of specific anti-parent CTLs was strictly contingent on co-culture of responder splenocytes with irradiated stimulator cells.

F_1 anti-parent cytotoxic responses are not limited to the particular strain combination of B6D2F_1 and B6[10,13–15]. CTL specificity was always determined by the strain of the parental cells—B6D2F_1 spleen cells generate anti-B6 or anti-DBA/2 CTLs depending on whether B6 or DBA/2 cells were the stimulators[10]. Similarly, (B6×C3H)F_1 CTLs specific for either B6 or C3H targets can be elicited by using B6 or C3H stimulator cells[14].

The generation of F_1 anti-parent CTLs occurred in conditions identical to those of the conventional, antigen-dependent, allogeneic mixed leukocyte reaction. As the F_1 mice were the offspring of crosses between inbred strain parents, F_1 anti-parent immune responsiveness could be attributed to the exclusive expression of certain alloantigens on homozygous parental cells. Such an interpretation contradicts the immunogenetic laws of transplantation and of immunity to histocompatibility antigens[16,17]. According to these empirical generalizations, the expression of membrane-bound alloantigens eliciting cell-mediated immunity is co-dominant and, thus, not limited to homozygous parental cells. Irradiation of parental

0028–0836/81/070559–05$01.00

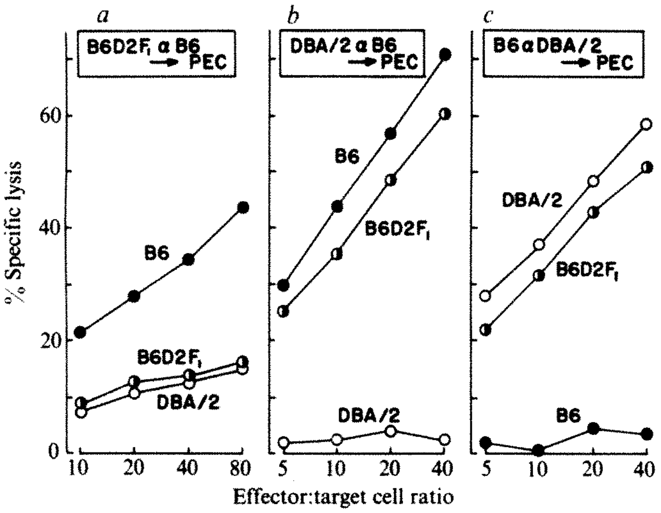


Fig. 1 Direct cytotoxicity of F_1 hybrid anti-parent (*a*) and allogeneic (*b*, *c*) CTLs induced *in vitro*. Primary 5-day mixed-splenocyte cultures of responder and irradiated (2,000 rad of ^{137}Cs γ rays) stimulator cells were set up as described[10,12,14]. The stimulator:responder ratio was 1:1 for F_1/parent and 1:10 for allogeneic cell mixtures; the number of responder cells was held constant at 3.5×10^6 per culture. Cultured spleen cells were tested for cytotoxic activity in 4-h chromium release assays on PECs stimulated by thioglycollate and prelabelled with radioactive ^{51}Cr at the indicated effector:target cell ratios[12,14]. The strains of responder and stimulator cell donors are given in the rectangle at the top of each panel; those of target PEC donors are indicated next to the lines relating cytolysis to the effector:target cell ratio. Per cent specific lysis was calculated according to the formula:

$$\% \text{ Specific lysis} = \frac{\text{experimental} - \text{spontaneous}}{\text{maximum} - \text{spontaneous}} \times 100$$

Spontaneous release was the radioactivity released by target cells in the presence of cultured but nonstimulated spleen cells. Spontaneous release was 15–20% of the maximum release, that is, the radioactivity released by target cells frozen and thawed four times. Experimental release was the radioactivity released by target cells in the presence of cultured, stimulated cells. Standard errors of the mean per cent specific lysis of four replicate samples were consistently below 10% of the mean values. Note that B6D2F$_1$ and DBA/2 PEC targets were lysed to a very limited extent by B6D2F$_1$ anti-B6 CTLs. In contrast, target cells of the same preparations were fully susceptible to lysis by allogeneic CTLs of appropriate specificity.

stimulator cells could have altered membrane structures so as to render them more immunogenic or to express new determinants, but this is unlikely as F_1 anti-parent CTLs can be elicited by nonirradiated, mitomycin C-treated stimulator cells (ref. 15 and our unpublished data) and the CTLs are not specific for irradiated target cells. Generation of autoreactive CTLs in F_1 anti-parent mixed cultures is an alternative explanation. The exclusive lysis of parental targets would then be attributed to the density or other property of membrane epitopes that may distinguish homozygous parental from heterozygous F_1 cells. Although this second interpretation of F_1 anti-parent cytotoxicity is not consistent with immunological theories postulating the elimination of autoreactive clones of lymphocytes at an early stage of development[18,19], the data below show it to be valid.

F_1 hybrid anti-parent CTLs are specific for H-2-controlled molecules

The genetic control of target determinants was investigated by replacing B6 targets (H-2^b) with PECs of various H-2 types (Table 1). If B6D2F$_1$ anti-B6 CTLs were specific for H-2 determinants, any PECs of H-2^b type should be lysed irrespective of the strain of origin, whereas PECs of H-2^a, H-2^d or H-2^k type should not be lysed even if donor mice shared the genetic backgrounds with strains B6 or C57BL/10 (abbreviated B10). The anti-parent CTLs lysed B6, B10 and C3H.B10 PECs, all H-2^b (group 1), but not PECs of strain A, B10.A (H-2^a), DBA/2, B10.D2 (H-2^d) and C3H (H-2^k) mice (group 2). PECs carrying the antigens of recombinant H-2 haplotypes involving chromosomal exchanges between H-2^b and H-2^a, or H-2^b and H-2^d, were lysed only when the b allele was present in the D region (group 3). The genetic background and the presence of the b allele in the K, I-A and I-B regions was irrelevant for target cell lysis as revealed by PECs of B10.A(5R) mice. PECs of F_1 hybrid mice with one copy of the original and one of the recombinant H-2 chromosomes were also tested as targets (groups 4 and 5)—D^b-homozygous PECs were the only F_1 target cells lysed. For all other PECs of F_1 hybrid mice, the values of specific lysis did not exceed background.

In conclusion, the lytic activity of B6D2F$_1$ anti-B6 CTLs was specific for D^b molecules, as indicated by the recombinant analysis of Table 1 and by additional studies with strains dissociating the control of target antigens from genes on both sides of H-2D, that is, from the S region on the centromeric and the Tla region on the telomeric sides of the chromosome[20]. The lack of reactivity of B6D2F$_1$ anti-B6 CTLs with cells of the H-2^a, H-2^d and H-2^k haplotypes suggests that determinants unique to the D^b molecules served as the target antigens. PECs of H-2^q, H-2^s and H-2^u types were also not lysed by F_1 CTLs raised against B6, and antigens of fetal calf serum did not take part in the lytic phase of this response[20].

F_1 hybrid anti-parent CTLs are autoreactive

The reactivity of B6D2F$_1$ anti-B6 CTLs generated in primary mixed spleen-cell cultures was further analysed by competitive inhibition experiments. The lysis of radiolabelled (hot) B6 PECs by F_1 CTLs was tested in the presence of added unlabelled (cold) PECs taken from mice of various inbred strains and F_1 crosses

Table 1 Genetics of target determinants for the lytic activity of B6D2F$_1$ anti-B6 CTLs

Group no.	Target cells Strain	H-2 haplotype K I SD ABJEC	% Specific lysis (effector:target cells = 40:1)
1	B6	bbbbbbbb	36.0
	B10	bbbbbbbb	32.4
	C3H.B10	bbbbbbbb	34.3
2	A	kkkkkddd	9.0
	B10.A	kkkkkddd	8.2
	DBA/2	dddddddd	5.8
	B10.D2	dddddddd	4.4
	C3H	kkkkkkkk	12.7
3	B10.A(4R)	kkbbbbbb	23.2
	B10.A(2R)	kkkkkddb	26.0
	HTG	dddddddb	38.0
	B10.A(5R)	bbbkkddd	8.4
4	B6×DBA/2	bbbbbbbb dddddddd	8.6
	B6×C3H	bbbbbbbb kkkkkkkk	11.1
5	B10×B10.A(2R)	bbbbbbbb kkkkkddd	
	B10.A×B10.A(2R)	kkkkkddd kkkkkddb	14.1
	B10.A×B10.A(5R)	kkkkkddd bbbkkddd	13.4

F_1 anti-parent CTLs were induced in mixed spleen-cell cultures and tested for cytolytic activity as indicated in the Fig. 1 legend. Target cells were radiolabelled PECs. Alleles at various regions of H-2 are given for the inbred strains used according to recent listings[49,50]. The data of this table are a composite of several experiments. Note that only H-2D^b-homozygous cells were effectively lysed (range of specific lysis 23–36%). All other targets were not effectively lysed: the range of 4–14% specific lysis is regarded as background. These low values varied between experiments and were dependent on the batch of fetal calf serum used during culture. Background lysis did not exhibit any identifiable specificity.

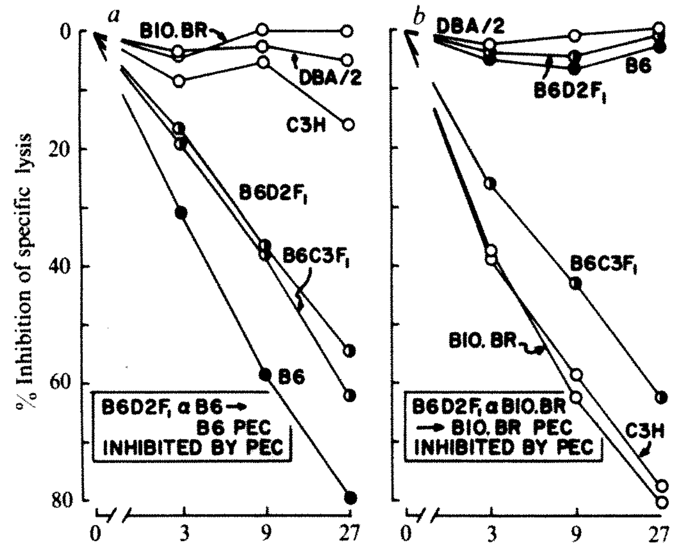


Fig. 2 Competitive inhibition of F_1 hybrid anti-parent (*a*) and F_1 anti-allogeneic (*b*) cell-mediated lympholysis by cold PECs. CTLs were raised in primary mixed spleen-cell cultures and assayed on PEC targets, as described in Fig. 1 legend. The strains of responder, stimulator and target cell donors are given in the rectangle of each panel. The effector : target cell ratio was 40 : 1 in F_1 anti-parent and 20 : 1 in F_1 anti-allogeneic cytotoxicity assays. In the absence of inhibitor PECs, specific lysis of radiolabelled parental B6 targets was 20.5%; that of allogeneic B10.BR(H-2^k) targets was 51.2%. Graded numbers of unlabelled inhibitor PECs were incubated with CTLs for 15–30 min before addition of radiolabelled target PECs. Thus, the number of inhibitor cells was the only variable in the assay mixture. The strains of inhibitor cell donors are given next to the line relating inhibition of lysis to the inhibitor : target cell ratio. Per cent inhibition of specific lysis was calculated according to the formula:

$$\%\ \text{Inhibition of specific lysis} = \left(1 - \frac{\%\ \text{specific lysis in the presence of inhibitors}}{\%\ \text{specific lysis in the absence of inhibitors}}\right) \times 100$$

Note that parental but not allogeneic target cell lysis was inhibited by $B6D2F_1$ PECs syngeneic with the effector cells in both tests. $B6C3F_1$ cells inhibited the lysis of parental targets because of the H-2^b haplotype, and that of allogeneic targets because of the H-2^k haplotype. See Table 1 for the H-2 haplotypes of inbred and F_1 mice used.

(Fig. 2). If the F_1 CTLs were specific for parental antigens not expressed on heterozygous cells, only H-2^b-homozygous, cold PECs should have been inhibitory. However, H-2^b-heterozygous $B6D2F_1$ and $B6C3F_1$, as well as homozygous B6, cold PECs were all able to compete with the hot B6 targets for the F_1 CTLs (Fig. 2*a*). PECs of C3H, B10.BR (H-2^k) and DBA/2 (H-2^d) mice, not bearing the D^b gene product, failed to inhibit target cell lysis, as expected. Heterozygous F_1 PECs were less efficient than homozygous B6 cells in inhibiting F_1 anti-parent cytolysis, presumably due to different gene dosage. The intrinsic ability of cold PECs to inhibit competitively and specifically the lytic activity of CTLs was verified by raising $B6D2F_1$ anti-allogeneic CTLs and conducting simultaneous inhibition assays for allogeneic cytolysis with given cold PEC preparations (Fig. 2*b*). The lytic activity of $B6D2F_1$ anti-allogeneic CTLs was unaffected by the presence of the same syngeneic, cold PECs that inhibited $B6D2F_1$ anti-parental cytotoxicity. Thus, F_1 anti-parent but not F_1 anti-allogeneic CTLs are autoreactive according to competitive syngeneic inhibition tests. A highly unusual feature of the autoreactive CTLs is the ability to bind specifically syngeneic PECs (Fig. 2) without apparently causing lysis (Fig. 1). It is this feature that operationally distinguishes autoreactive CTLs from H-2-restricted effectors specific for foreign antigens. These CTLs, though unable to kill syngeneic targets, are regarded as autoreactive because they specifically bind to target H-2 molecules expressed on their own surface as well as on their precursors.

Inhibition of F_1 anti-parent CTLs was not contingent on the use of PECs as inhibitor cells. Identical results were obtained in experiments in which H-2^b-homozygous, H-2$^{b/d}$-heterozygous and H-2^d-homozygous embryonic fibroblasts served as inhibitor cells instead of PECs (Fig. 3). This experiment was critical for concluding that F_1 anti-parent CTLs are autoreactive. Fibroblasts, unlike lymphocytes, lack receptors for target-cell alloantigens and therefore could not inhibit the lysis of radiolabelled PECs by binding to and masking antigenic sites. Blocking of cytolysis, as opposed to competition for CTLs, was not excluded in experiments with cold PECs which contain assorted lymphoid and nonlymphoid cell types. In addition, fibroblasts are unlikely to express idiotypic determinants and, thus, the data are not consistent with the possibility that F_1 responder lymphocytes engaged in anti-idiotypic cytotoxic responses restricted by H-2.

Genetic mapping of self determinants recognized by F_1 anti-parent CTLs

PECs from inbred and F_1 hybrid mice bearing H-2^b recombinant haplotypes (used in the genetic mapping of target determinants for direct lysis (Table 1)) were tested for the ability to inhibit competitively lysis of parental B6 targets by $B6D2F_1$ anti-B6 CTLs (Fig. 4). All homozygous and heterozygous cells that were D^b-positive inhibited cytolysis, whereas cells with other alleles in the D region, or with the b allele in the K, I-A and I-B regions, failed to inhibit (Fig. 4*a*). The inbred and F_1 cells of HTG, B10.A(2R) and B10.A(4R) origin or parentage were critical in demonstrating that one or two copies of D^b genes were necessary and sufficient for inhibition (see Table 1 for haplotype constitution). All of the critical cells used in this experiment could inhibit allogeneic CTLs raised against either D^b or K^b antigens (Fig. 4*b*, *c*). These controls persuaded us that the specificity of F_1 anti-parent CTLs, rather than possible variations in PEC preparations, determined the outcome of the assay. As the genetic mapping of target antigens for direct lysis (Table 1) agrees fully with that for competitive binding of homozygous

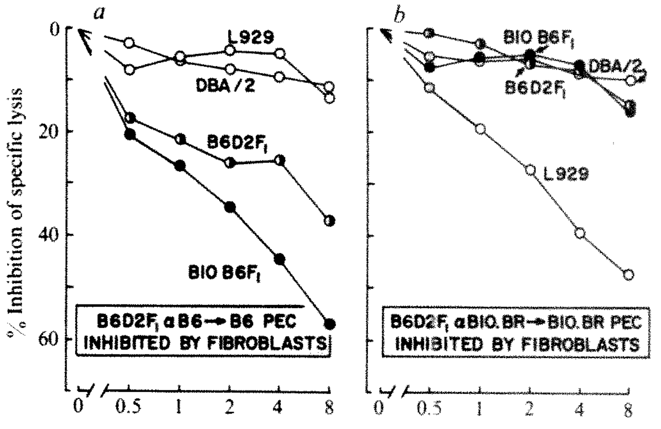


Fig. 3 Competitive inhibition of F_1 hybrid anti-parent (*a*) and F_1 anti-allogeneic (*b*) cell-mediated lympholysis by cold embryonic fibroblasts. CTLs were raised in primary mixed spleen-cell cultures and assayed on PEC targets as described in Figs 1 and 2 legends. The strains of responder, stimulator and target-cell donors are given in the rectangle of each panel. In the absence of inhibitor cells, specific lysis of radiolabelled parental B6 targets was 24.9%; that of allogeneic B10.BR(H-2^k) targets was 66.1%. Inhibition tests of cytolysis were carried out as described in Fig. 2 legend except for replacing PECs with cultured embryonic fibroblasts or transformed L929 fibroblasts. The strains of inhibitor cell donors are given next to the lines relating inhibition of lysis to the inhibitor : target cell ratio. $B10B6F_1$ stands for $(B10 \times B6)F_1$ hybrids (homozygous H-2^b). Primary monolayer cultures of fibroblasts were set up from near-term embryos and maintained for four to six serial passages before use. L929 fibroblasts are of C3H origin (H-2^k). Note the similarity between results of inhibition tests with PECs (Fig. 2) and with fibroblasts.

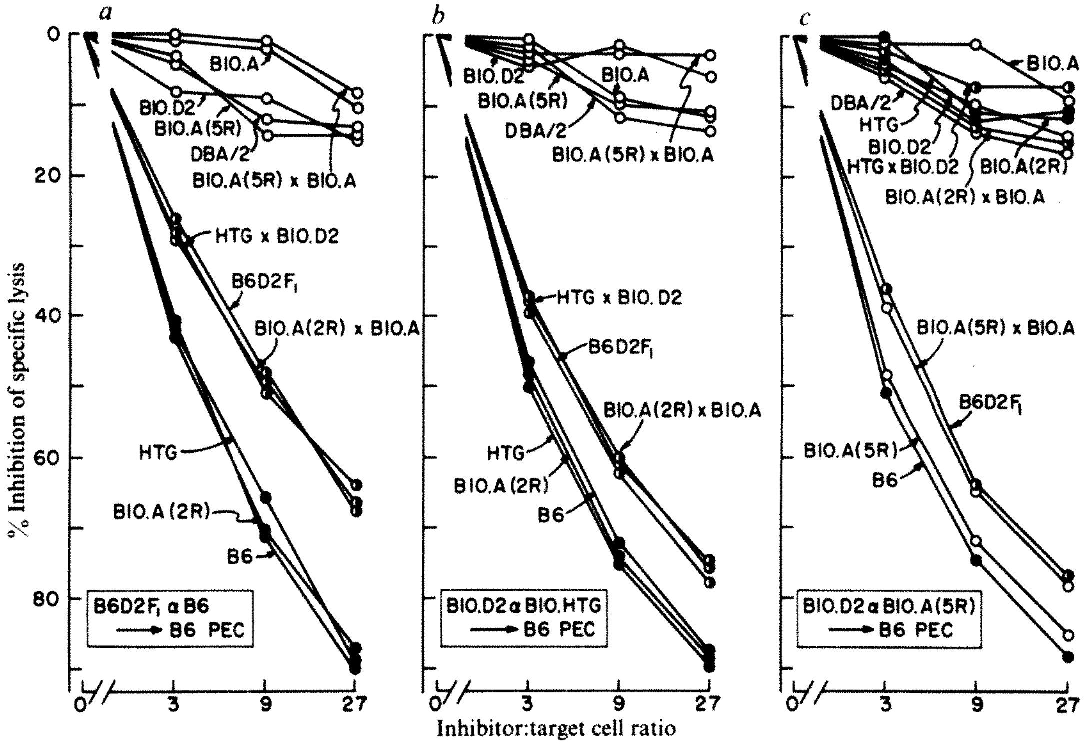


Fig. 4 Genetics of target determinants for the binding activity of B6D2F$_1$ anti-B6 CTLs in competitive inhibition tests. F$_1$ anti-parent and allogeneic CTLs were raised in primary mixed spleen-cell cultures and assayed on PEC targets, as described in Figs 1 and 2 legends. The strains of responder-, stimulator- and target-cell donors are given in the rectangle of each panel. In the absence of inhibition, specific lysis of radiolabelled parental B6 targets by F$_1$ CTLs was 23.7% (*a*); specific lysis of the same targets by allogeneic anti-H-2D^b (*b*) and anti-H-2K^b (*c*) CTLs was 40.2% and 33.1%, respectively. Inhibition tests of cytolysis were carried out as described in Fig. 2 legend. The strains of inhibitor PEC donors are given next to the inhibition lines; H-2 haplotypes of the mice used are indicated in Table 1. Note that one copy of the D^b allele was necessary and sufficient for [B10.A(2R)×B10.A]F$_1$ and (HTG×B10.D2)F$_1$ PECs to inhibit cytolysis by B6D2F$_1$ CTLs (*a*). Similar results were obtained with [B10.A(4R)×B10.A]F$_1$ and B10.A(4R) PECs (not shown because of crowding). Note also the verification of the competence and specificity of D^b+K^b−, D^b+K^b+, D^b−K^b+ and D^b−K^b− PECs for inhibiting allogeneic cytolysis directed against either D^b (*b*) or K^b (*c*) alloantigens.

and heterozygous PECs (Fig. 4), we conclude that the reactivity of B6D2F$_1$ anti-B6 CTLs is directed to self antigens controlled by H-2D^b. Studies with mixed splenocyte cultures of other F$_1$/parental strain combinations indicate that this type of autoreactivity is not exceptional—F$_1$ CTLs raised against parental H-2^d or H-2^k cells are specific for self antigens controlled by H-2D^d, H-2L^d (ref. 21) and H-2K^k (ref. 14), respectively. Such CTLs also bind to homozygous and heterozygous PECs, but only lyse the former (I. N., J. F. Warner and G.C. unpublished data). To our knowledge, there is no precedent for lysis-negative, inhibition-positive target cells in specific cell-mediated immune responses.

Requirements for induction of anti-self H-2 immune responses

This and other publications[9-15,20-22] and our unpublished studies show that effector T lymphocytes specific for self H-2 products can readily and consistently be generated *in vitro* from spleen cells of conventional, untreated, adult F$_1$ hybrid mice. The target structures for such CTLs are expressed on normal and transformed lymphoid cells[9-15], fibroblasts[22] and peritoneal exudate cells[12,14]. Most important, nonstimulated peritoneal cells freshly explanted and maintained in medium free of fetal calf serum served as targets, indicating that normal cells do express the target structures for autoreactive CTLs[20].

The major genetic requirement was the H-2 heterozygosity of responder cells to be cultured with homozygous stimulator cells. In the conditions used, autoreactive CTLs were not obtained in primary cultures of H-2-homozygous splenocytes with and without irradiated syngeneic or allogeneic cells. Autoreactive F$_1$ CTLs were specific for molecules controlled by one end of the non-recombinant H-2 complex even though heterozygosity extended to both ends[14,15,20,21]. F$_1$ spleen cells are competent, however, to generate more than one major population of CTLs—B6D2F$_1$ anti-DBA/2 CTLs reactive with D-end products include at least two independent subpopulations, one specific for D^d and the other for L^d molecules[21]. Moreover, the stimulation of F$_1$ splenocytes with homozygous cells bearing an appropriate H-2 recombinant haplotype may result in the independent generation of CTL clones specific for products of both ends of H-2, for example, D^b and K^k (ref. 14).

The particular conditions in which autoreactive CTLs were raised may account for, first, the requirement of responder-cell heterozygosity, and second, the CTL properties of strict stimulator-directed specificity and of inability to lyse heterozygous target cells. Major phenotypic differences between the homozygous and heterozygous cells are due to the effect of gene dosage. Although H-2 gene products are co-dominantly expressed on heterozygous cells[23], the products of both alleles may not necessarily be expressed in equivalent amount or fashion[24]. Moreover, interallelic and intergenic 'hybrid or interaction determinants' exist for murine I-region products[25,26] and for erythrocyte antigens of rabbits[27] and pigeons[28]. It is, therefore, conceivable that allelic interactions at the genetic and phenotypic levels result in quantitative and/or qualitative differences in expression of H-2 molecules on the surface of heterozygous and homozygous cells. Such differences could explain why homozygous cells provide optimal immunogenic stimuli for the induction of autoreactive F$_1$ CTLs, and why the binding of CTLs to heterozygous targets fails to trigger lytic events. However, noncytolytic F$_1$ cells of the effector preparation may exert autoinhibition and the results would be the same—lysis of heterozygous targets would be inhibited more effectively than that of homozygous targets.

It has been shown previously that irradiated parental spleen cells depleted of T lymphocytes[12] or of phagocytic cells adherent to plastic or Sephadex G-10 (ref. 29) cannot function as stimulators of F$_1$ anti-parent responses. Parental phagocytic-adherent cells expressing I-region antigens are the most effective antigen-presenting cells in this system (ref. 29 and unpublished data of W. W. Freimuth and G.C.). Parental T lymphocytes, which by themselves stimulate rather weakly, may provide a triggering signal through back stimulation and release of lymphokines[12]. The effectiveness of purified, phagocytic-adherent stimulator cells may be attributed to presentation of self H-2 molecules to CTL precursors, coupled with the induction of syngeneic mixed-leukocyte reactions[30-32] and release of cytokines or other factors. Thus the generation of autoreactive CTLs involves at least two steps: (1) activation of F_1 CTL precursors by nonspecific diffusible mediators, and (2) selection and differentiation of CTL clones driven by self immunogen.

Autochthonous or spontaneous *in vitro* generation of anti-self H–2 CTLs by lymphoid cells of inbred or F_1 mice has been described by several investigators[32–36]. Deliberate stimulation was either not applied or was provided for by irradiated syngeneic spleen cells. In the context of the suggested model, allogeneic effect factor[32,33] and fetal calf serum[34–36] may have provided directly or indirectly the activating stimuli. Autologous or syngeneic cells present in the culture may have been the source of immunogen. In the absence of potent helper factors, or in conditions less favourable than those of mixed F_1/parent spleen-cell cultures, the generation of autoreactive CTLs is likely to be erratic and part of a broader, polyclonal lymphocyte activation[37–39]; such CTLs may be restricted by, rather than directed against, self H–2. In an interesting model of spontaneous generation of CTLs[36], the age of F_1 mice donating lymphoid cells was a critical variable, and the expression of target determinants on parental and syngeneic F_1 lymphoblasts was transient after stimulation with mitogen. Endogenous viral antigens that become expressed on stimulated lymphocytes[40–42] may have been recognized by these CTLs in the context of self H–2 determinants.

The induction of F_1 anti-parent CTLs has its closest parallel in the production of anti-self H–2 antibodies by F_1 hybrid mice immunized with cells of a parental Abelson virus-induced lymphoma[43]. The antibodies were specific for H–2D and H–2K antigens expressed on normal and neoplastic cells of H–2 homozygous and heterozygous mice. The successful induction of autoreactive B-lymphocyte effectors specific for self H–2 antigens apparently depended on the heterozygosity of responder mice, homozygosity of immunizing cells and probably on the diffusible mediators that virus-infected lymphoma cells release or induce. Hybrid resistance to parental grafts of normal or malignant haematopoietic cells[9,44–46] may represent yet another response to self H–2 determinants mediated by effectors of the natural killer (NK) type[47,48]. Like the autoreactive T- and B-lymphocyte responses, this F_1 host resistance involves recognition of, and reactivity against, the products of H–2-associated genes of homozygous parental target cells. However, no evidence has been obtained for cytotoxic or other activity (for example, cytostatic) of NK-like effectors against syngeneic targets.

We thank Vita Milisauskas for cultured embryonic fibroblasts. K.N. was on leave from the Department of Medicine and Physical Therapy, University of Tokyo, Japan. This work was supported by NIH grants AM-13969 and CA-12844, and American Cancer Society grant IM-212.

Received 11 September; accepted 4 November 1980.

1. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27,** 52–177 (1979).
2. Shearer, G. M. & Schmitt-Verhulst, A. -M. *Adv. Immun.* **25,** 55–91 (1977).
3. Wagner, H., Starzinski-Powitz, A., Jung, H. & Röllinghoff, M. *J. Immun.* **119,** 1365–1368 (1977).
4. Billings, P., Burakoff, S., Dorf, M. E. & Benacerraf, B. *J. exp. Med.* **146,** 623–628 (1977).
5. Bevan, M. J. *J. exp. Med.* **142,** 1349–1364 (1975).
6. Schwartz, R. H., Yano, A. & Paul, W. E. *Immun. Rev.* **40,** 153–180 (1978).
7. Katz, D. H. in *Lymphocyte Differentiation, Recognition and Regulation* 550–594 (Academic, New York, 1977).
8. Dausset, J. & Contú, L. *Hum. Immun.* **1,** 5–17 (1980).
9. Shearer, G. M. & Cudkowicz, G. *Science* **190,** 890–892 (1975).
10. Shearer, G. M., Garbarino, C. A. & Cudkowicz, G. *J. Immun.* **117,** 754–759 (1976).
11. Shearer, G. M. *et al. Cold Spring Harb. Symp. quant. Biol.* **41,** 511–518 (1977).
12. Nakamura, I. & Cudkowicz, G. *Eur. J. Immun.* **9,** 371–379 (1979).
13. Schmitt-Verhulst, A. -M. & Zatz, M. M. *J. Immun.* **118,** 330–333 (1977).
14. Warner, J. F. & Cudkowicz, G. *J. Immun.* **122,** 575–581 (1979).
15. Ishikawa, H. & Dutton, R. W. *J. Immun.* **122,** 529–536 (1979).
16. Snell, G. D. & Stimpfling, J. H. in *Biology of the Laboratory Mouse* (ed. Green, E. L.) 457–491 (McGraw-Hill, New York, 1966).
17. Snell, G. D. *Adv. Genet.* **20,** 291–355 (1979).
18. Burnet, F. M. *The Clonal Selection Theory of Acquired Immunity* (Cambridge University Press, London, 1959).
19. Jerne, N. K. *Eur. J. Immun.* **1,** 1–9 (1971).
20. Nakano, K., Nakamura, I. & Cudkowicz, G. *J. Immun.* (in the press).
21. Nakamura, I., Nakano, K. & Cudkowicz, G. *Immunogenetics* **9,** 317–318 (1979).
22. Ilfeld, D., Carnaud, C. & Klein, E. *Immunogenetics* **2,** 231–240 (1975).
23. Cullen, S. E., Schwartz, B. D., Nathenson, S. G. & Cherry, M. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1394–1397 (1972).
24. O'Neill, H. C. & Blanden, R. V. *J. exp. Med.* **149,** 724–731 (1979).
25. Fathman, C. G., Watanabe, T. & Augustin, A. *J. Immun.* **121,** 259–264 (1978).
26. Lafuse, W. P., McCormick, J. F. & David, C. S. *J. exp. Med.* **151,** 709–715 (1980).
27. Cohen, C. *Science* **123,** 935–936 (1955); *J. Immun.* **84,** 501–506 (1960).
28. Irwin, M. R. & Cole, L. J. *J. exp. Zool.* **73,** 85–108 (1936); *J. exp. Zool.* **73,** 309–318 (1936).
29. Cudkowicz, G., Yung, Y. P. & Freimuth, W. W. in *Natural and Induced Cell-Mediated Cytotoxicity* (eds Riethmüller, G., Wernet, P. & Cudkowicz, G.) 209–220 (Academic, New York, 1979).
30. Howe, M. L., Goldstein, A. L. & Battisto, J. R. *Proc. natn. Acad. Sci. U.S.A.* **67,** 613–619 (1970).
31. Ponzio, N. M. *Cell Immun.* **49,** 266–282 (1980).
32. Altman, A. & Katz, D. H. *Immun. Rev.* **51,** 3–34 (1980); *J. Immun.* **125,** 1536–1543 (1980).
33. Altman, A., Cardenas, J. M., Bechtold, T. E. & Katz, D. H. *J. Immun.* **124,** 114–120 (1980).
34. Peck, A. B., Andersson, L. C. & Wigzell, H. *J. exp. Med.* **145,** 802–818 (1977).
35. Ching, L. M., Marbrook, J. & Walker, K. Z. *Cell. Immun.* **31,** 284–301 (1977).
36. Ishikawa, H. & Dutton, R. W. *J. Immun.* **125,** 656–662 (1980).
37. Shustik, C., Cohen, I. R., Schwartz, R. S., Latham-Griffin, E. & Waksal, S. D. *Nature* **263,** 699–701 (1976).
38. Burton, R. C., Chism, S. E. & Warner, N. L. *J. Immun.* **119,** 1329–1339 (1977).
39. Golstein, P., Luciani, M. -F., Wagner, H. & Röllinghoff, M. *J. Immun.* **121,** 2533–2538 (1978).
40. Wecker, E., Schimpl, A. & Hünig, T. *Nature* **269,** 598–600 (1977).
41. Moroni, C., Schumann, G., Robert-Guroff, M., Suter, E. R. & Martin, D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 535–538 (1975).
42. Phillips, S. M. *et al. Cell. Immun.* **15,** 169–179 (1975).
43. Risser, R. & Grunwald, D. J. *Nature* **289,** 563–568 (1981).
44. McCulloch, E. A. & Till, J. E. *J. Cell. comp. Physiol.* **61,** 301–308 (1963).
45. Cudkowicz, G. & Stimpfling, J. H. *Immunology* **7,** 291–306 (1964); *Science* **147,** 1056 (1965).
46. Cudkowicz, G. in *21st A. M. D. Anderson Symp. on Fundamental Cancer Res.* 661–691 (Williams & Wilkins, Baltimore, Maryland, 1968); in *Natural Resistance Systems Against Foreign Cells, Tumors and Microbes* (eds Cudkowicz, G., Landy, M. & Shearer, G. M.) 3–30 (Academic, New York, 1978).
47. Cudkowicz, G. & Bennett, M. *J. exp. Med.* **134,** 1513–1528 (1971).
48. Kiessling, R. *et al. Eur. J. Immun.* **7,** 655–663 (1977).
49. *Inbred and Genetically Defined Strains of Laboratory Animals* Part 1 (eds Altman, P. L. & Katz, D. D.) 123–127 (Federation of American Societies for Experimental Biology, Bethesda, Maryland, 1979).
50. Klein, J., Flaherty, L., VandeBerg, J. L. & Shreffler, D. C. *Immunogenetics* **6,** 489–512 (1978).

Production of anti-self H–2 antibodies by hybrid mice immune to a viral tumour

Rex Risser & David J. Grunwald

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706

During the course of immunization of hybrid mice (genotype $H\text{–}2^{b/k}$*) with a parental Abelson virus-induced lymphoma (genotype* $H\text{–}2^{b/b}$*), antibodies were produced to the* $H\text{–}2K^b$ *or* $H\text{–}2D^b$ *antigens of the immunizing cells. Such 'anti-self H–2' antibodies demonstrate the existence of autoreactive B-cell clones in hybrid mice, and pose intriguing questions for the nature of self-tolerance.*

THE success or failure of tissue grafts among inbred mice has been found to conform to certain empirically derived laws, one being that grafts from an inbred homozygous parent are accepted by hybrid offspring made between the donor strain and another inbred strain[1,2]. A corollary of this law is that hybrids do not generate cell-mediated or humoral immune responses to the parental graft, that is, they are tolerant to grafts of parental tissues[3–5]. The major histocompatibility complex (MHC) of the

0028–0836/81/070563–06$01.00

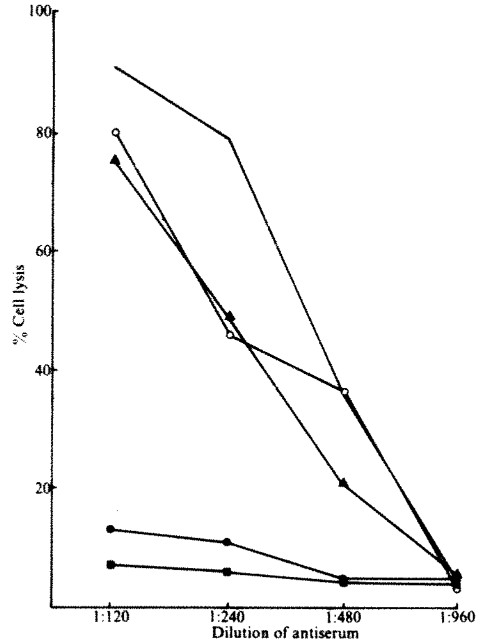


Fig. 1 Absorption of (C3H × B6)F_1 anti-B6T1 (S41 no. 2) antiserum with normal spleen cells from *H-2* congeneic mice. Cytotoxicity tests were performed on B6T1 cells as described[15] using preselected rabbit serum (dilution 1:7) as a complement source. Absorption experiments were performed as described elsewhere[15] by incubating 5×10^7 nucleated spleen cells with 120 μl of antiserum (dilution 1:120). Cells used for absorption: —, none; ●, B6 ($H\text{-}2^b$); ■, B10 ($H\text{-}2^b$); ○, B10.BR ($H\text{-}2^k$), and ▲, B6.C-H-2^d ($H\text{-}2^d$). Absorption with $H\text{-}2^b$ cells removed cytotoxic antibodies, whereas absorption with congeneic $H\text{-}2^d$ or $H\text{-}2^k$ cells did not.

mouse, *H-2*, controls these cell interactions, partly by specifying the genetically dominant expression of polymorphic transplantation antigens at the surfaces of all somatic cells[1,2,6].

Exceptions to the law of hybrid acceptance of parental grafts were noted in early tumour transplantation studies[7,8] and have also been observed in spleen and bone marrow cell transfers in certain strain combinations[9,10]. In the latter case it has been proposed that these exceptions reflect the expression of antigens by homozygous but not heterozygous mice[10]. A locus for such haematopoietic histocompatibility or Hh antigens displayed on parental cells and recognized by F_1 mice has been mapped to chromosome 17 of the mouse in the *H-2D* region[11]. The search for *in vitro* correlates of hybrid histocompatibility led to the discovery of induced and spontaneous F_1 anti-parent cytotoxic T-lymphocyte responses[12], the determinants of which also map to *H-2D* or, in some instances, to *H-2K*[13]. Because the recognition of a variety of antigens by cytotoxic T lymphocytes has been shown to require *H-2K* or *H-2D* identity on target and effector cells[14], it is unclear whether these F_1 anti-parent cytotoxic T-lymphocyte responses reflect responses to some additional antigen (perhaps Hh or viral) whose recognition is *H-2* restricted or true anti-MHC autoreactivity.

During the course of immunizations of various hybrid mice with a parental transplanted lymphoma induced by the Abelson murine leukaemia virus, we observed a serological correlate of hybrid rejection of parental tumour cells. Here we characterize this serological response and demonstrate that it is directed to the H-2K and H-2D molecules of the parental cells, although such products can also be found on cells of the responding hybrid.

A high titre of cytotoxic antibodies was produced when [C3H/HeJ × C57BL/6(B6)]F_1 mice were hyperimmunized with the transplanted B6 Abelson virus lymphoma B6T1[15]. Antiserum from one such hyperimmune mouse (S41 no. 2) was cytotoxic for the immunizing cell with a mid-point titre of 1:320. However, when antiviral antibodies were removed from the antiserum by absorption with tissue culture cells (non-$H\text{-}2^b$) producing virus, cytotoxic antibodies remained that reacted with the immunizing lymphoma cell. When the serum was absorbed a second time with bone marrow cells from each of the seven BALB/c × B6 recombinant inbred strains of Bailey[16], cells from strains carrying the $H\text{-}2^b$ allele of B6 removed cytotoxic reactivity for B6T1 but cells from the strains carrying $H\text{-}2^d$ of the BALB/c allele did not (data not shown). These results suggested that at least one antibody present in antiserum from mouse S41 no. 2 detected a cell-surface antigen on B6T1 that was shared with normal spleen cells carrying the $H\text{-}2^b$ allele. To confirm that such antibodies detected antigens controlled by the $H\text{-}2^b$ haplotype, unabsorbed antiserum was diluted 1:120 then absorbed once with spleen cells from uninfected B6, C57BL/10 (B10), congeneic B6.C-H-2^d or congeneic B10.BR ($H\text{-}2^{k/k}$) mice and tested in cytotoxic tests with B6T1 target cells. The results of this experiment, shown in Fig. 1, demonstrated that the major antigens detected on B6T1 cells in cytotoxic tests with this antiserum were also expressed on normal spleen cells of B6 or B10 mice but not on spleen cells of $H\text{-}2^d$ or $H\text{-}2^k$ congeneic partner strains.

If such antibodies were directed to the H-2 determinants expressed on normal cells, then they would be expected to lyse normal spleen cells from H-2^b-carrying mice. This antiserum lysed 95% of spleen cells from B6, B10, 129 and (C3H × B6)F_1 mice but did not lyse spleen cells from C3H/HeJ, B10.BR, B10.A or B6.C-H-2^d mice, a result that clearly demonstrated the reactivity of this serum for determinants specified by the $H\text{-}2^b$ haplotype[17] (Fig. 2).

Response of hybrid mice to immunization with B6T1

The unusual nature of the reactivities found in the preceding antiserum prompted us to examine several other hybrid mice for production of anti-self H-2 antibodies. To do this we screened for antibodies that lyse 95% of spleen cells from B6 or B10 mice but show low (0–25%) lysis of spleen cells from B6.C-H-2^d or B10.D2n($H\text{-}2^d$) mice in direct cytotoxic tests with these cells. The results of this survey (Table 1) indicated that several hybrid mice which survived hyperimmunization produced cytotoxic

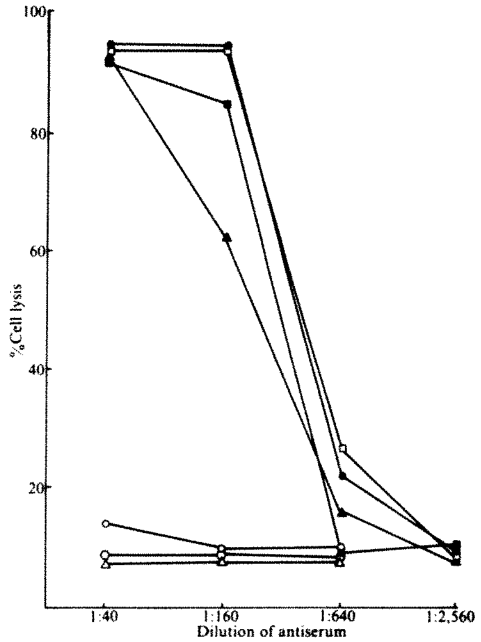


Fig. 2 Direct cytotoxic tests with (C3H × B6)F_1 anti-B6T1 (S41 no. 2) antiserum were performed as described[15] on spleen cells of: ●, B6; □, B10; ▲, 129 ($H\text{-}2^b$); ■, (C3H × B6)F_1 ($H\text{-}2^{k/b}$); ○, B10.BR; △, B6.C-H-2^d; and ○, C3H ($H\text{-}2^k$) mice using preselected rabbit serum (dilution 1:15) as a complement source. Normal spleen cells of mice bearing only the $H\text{-}2^b$ haplotype were lysed by this antiserum.

Table 1 Immunization and antibody response of hybrid mice

Mouse	Immunogen	No. mice immunized	No. surviving 10^8 cells	No. producing αH–2^b/No. screened
(C3H × B6)F_1 10–13 months	—	—	—	0/11
(BALB/c × B6)F_1 15 months	—	—	—	0/6
(C3H × B6)F_1	B6 bone marrow	10	10	0/10
B6	B6T1	113	15	0/15
(C3H × B6)F_1	B6T1	16	5	3/5
(C57BR × B6)F_1	B6T1	11	3	1/3
(BALB/c × B6)F_1	B6T1	15	2	1/2
(SEA × B6)F_1	B6T1	11	1	0/1

Female mice 8–12 weeks old were inoculated subcutaneously with 10^2 B6T1 cells, and tumours that developed were allowed to regress before a second intraperitoneal challenge of 10^2 cells was given. In alternate weeks mice were immunized with 10-fold increasing dose of cells until a plateau of 10^8 cells per mouse was achieved, and this immunization dose was maintained for the life of the animal. The initial immunization with bone marrow cells was 10^6 cells. Mice were bled individually 10 days after immunization, and antiserum, diluted 1 : 20, was screened for cytotoxic activity on B6 and B6.C–H–2^d, or B10 and B10.D2n spleen cells. Mice were considered positive for H–2^b antibody if their sera lysed 95% of spleen cells from the H–2^b strain. No antiserum sample lysed >30% of spleen cells from the respective congeneic (B6.C–H–2^d or B10.D2n) strains.

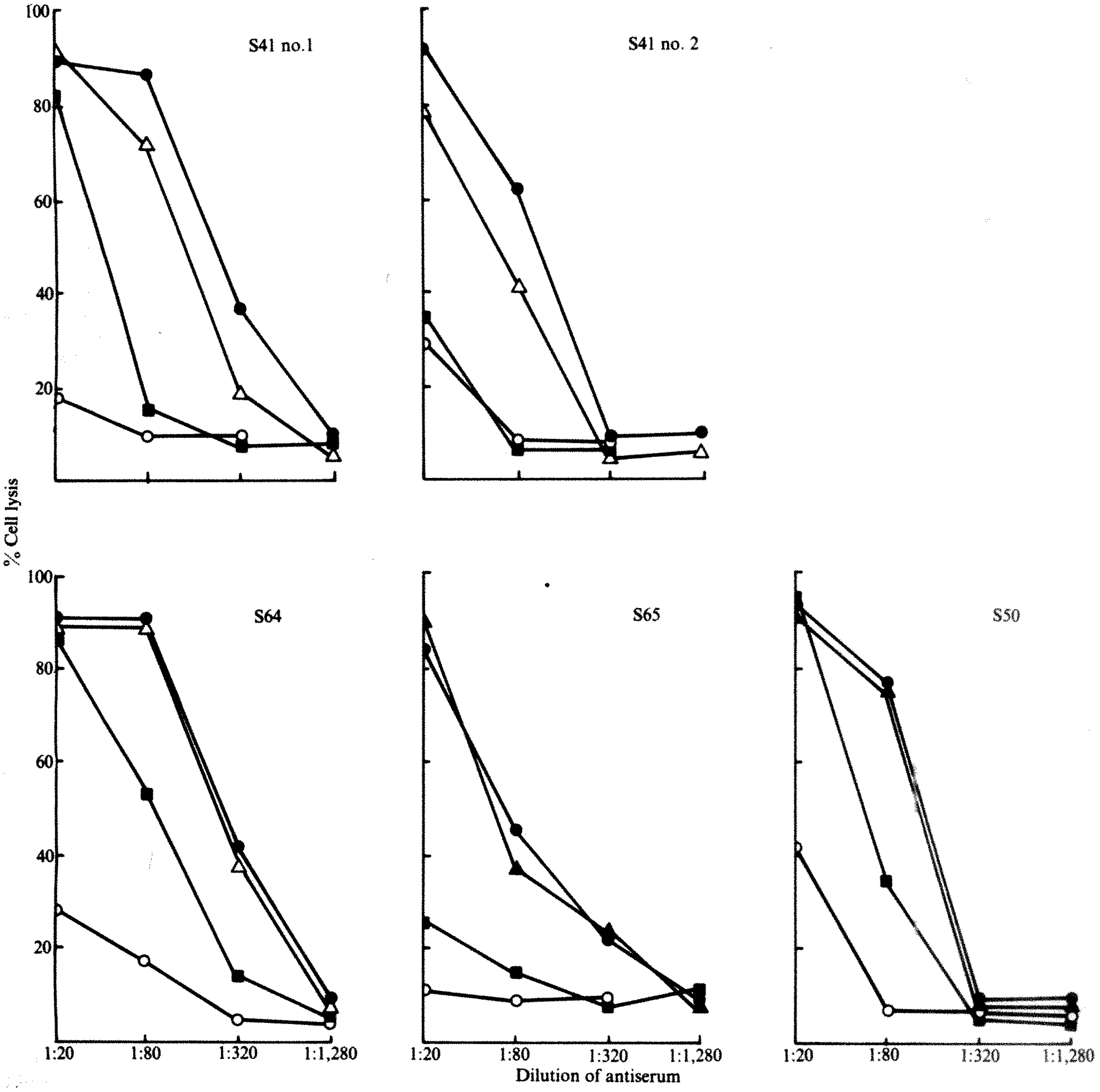


Fig. 3 Genetic mapping of antigens recognized by anti-self H–2 antisera. Antisera were assayed in direct cytotoxic tests on: ●, B10; △ and ▲, B10.A(5R) (*H–2K^b^, D^d^*); ■, B10.A(2R) (*H–2K^k^, D^b^*); and ○, B10.A (*H–2K^k^, D^d^*) cells as described elsewhere[15] using preselected rabbit serum (dilution 1 : 15) as a complement source. All five hybrid sera which contained anti-H–2^b antibodies recognized H–$2K^b$ antigens, and three of these sera (S41 no. 1, S64, S50) also recognized H–$2D^b$ antigens.

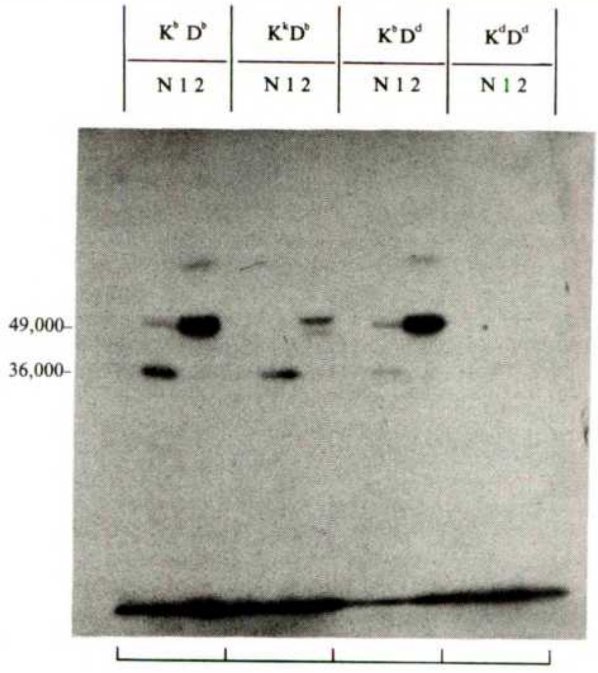


Fig. 4 Reactivity of antisera from hybrid mice with parental H–2K^b and/or H–2D^b molecules. 3×10^7 spleen cells from B10, B10.A(2R), B10.A(5R) and B10.D2n mice in 1 ml phosphate-buffered saline (PBS) were radiolabelled with 2 mCi Na^{125}I by the lactoperoxidase method[31,32]. Cells were washed with PBS, resuspended in 0.8 ml hypotonic buffer (10 mM Tris-HCl, *p*H 7.4, 10 mM NaCl, 1.5 mM $MgCl_2$, 1% Trasylol [protease inhibitor, FBA]) and lysed by the addition of 0.2 ml lysis buffer (50 mM $NaHPO_4$, *p*H 7.2, 4.5% NaCl, 5% Triton X-100, 2.5% sodium deoxycholate, 0.5% SDS)[33]. Lysates were centrifuged at 1,500*g* for 20 min to remove nuclei. Each lysate was then mixed with 3 ml of a 10% suspension of formalin-fixed *Staphylococcus aureus* bacteria (Cowan I strain)[34] and centrifuged at 120,000*g* for 90 min. This procedure removed all iodinated immunoglobulin molecules capable of binding the protein A-bearing bacteria. 10 μl of each antiserum was incubated with 2×10^6 c.p.m. of each lysate and immune complexes were collected by the addition of 200 μl of the *S. aureus* suspension. The immunoprecipitates were washed four times in RIPA buffer[35] (10 mM Tris-HCl, *p*H 7.2, 150 mM NaCl, 1% sodium deoxycholate, 1% Triton X-100, 0.1% SDS, 1% Trasylol), solubilized by heating to 100 °C for 5 min in sample buffer (0.0625 M Tris-HCl, *p*H 6.8, 10% glycerol, 5% 2-mercaptoethanol, 3% SDS, 0.001% bromophenol blue) and electrophoresed through a discontinuous 10% SDS-polyacrylamide slab gel according to the method of Laemmli[36]. For visualization of protein markers, gels were fixed and stained in 0.05% Coomassie brilliant blue in a solution containing ethanol/acetic acid/water (5 : 1 : 5) and destained in 7.5% acetic acid/5% ethanol. Dried gels were exposed to Kodak X-Omat R film in the presence of an intensifying screen. K^bD^b, K^kD^b, K^bD^d and K^dD^d refer to the *H–2* genotypes of the spleen cells from the B10, B10.A(2R), B10.A(5R) and B10.D2n strains, respectively. Tracks labelled N, 1 and 2 refer to the antisera used: N, normal mouse serum; 1, (C3H × B6)F_1 anti-B6T1, S41 no. 2; and 2, (BALB/c × B6)F_1 anti-B6T1, S50.

anti-H–2^b antibodies. As the data in Table 1 indicate, most of the mice challenged with this tumour do not survive, thus insufficient numbers were available to determine whether the production of anti-self H–2 antibodies was detrimental to the health of the animals. We can say, however, that the five hybrid mice which produced anti-H–2^b antibodies survived at least six challenges with 10^8 B6T1 cells, and one mouse, a (C57BR × B6)F_1, lived 15 months (20 challenges of 10^8 B6T1 cells) after its initial anti-H–2^b response was detected. This mouse (S65) was killed because it seemed to be lymphomatous. Autopsy revealed a massive lymphoma involving the mesenteric lymph node, liver and thymus, and histological examination also revealed glomerulonephritis. Washed cells from the tumour were lysed in cytotoxicity tests with conventional anti-H–2^b and anti-H–2^k antisera as well as the anti-self H–2 antiserum of that mouse, thus the lymphoma was of recipient and not immunizing cell origin. It remains to be determined whether the tumour, which was of recipient genotype, arose as a consequence of virus infection or autoimmune reactions.

Sera from 11 aged non-immune (C3H × B6)F_1 or six aged non-immune (BALB × B6)F_1 mice did not show appreciable reactivity (≤15% lysis) for B6 spleen cells, nor did antisera from 10 (C3H × B6)F_1 mice hyperimmunized with B6 bone marrow cells show reactivity for H–2^b spleen cells. Furthermore, antisera from 15 B6 mice hyperimmunized with the same B6T1 lymphoma cells did not have significant cytotoxic activity for B6 spleen cells, although all of these antisera contained antibody cytotoxic for B6 Moloney murine leukaemia virus (MuLV) leukaemias (recognition of the Friend–Moloney–Rauscher antigen[18]), and several contained antibody specifically cytotoxic for B6 Abelson lymphomas (recognition of the Abelson lymphoma antigen[15]).

Genetic mapping of determinants recognized by anti-self H–2 antibodies

Antisera from each of the five mice that produced anti-H–2^b antibodies were tested for direct cytotoxic activity on spleen cells of B10, B10.A, B10.A(2R) and B10.A(5R) mice (Fig. 3). All antisera reacted well with B10 and B10.A(5R) cells as >95% of spleen cells were lysed with titration mid-points of 1 : 160–1 : 640. Two antisera also reacted well with B10.A(2R) cells, and one antiserum (S64) reacted weakly with B10.A(2R) (95% lysis with a titration mid-point of 1 : 40). None of the antisera lysed more than 30% of spleen cells from B10.A. These results indicate that antigens controlled by the *H–2K–IA–IB* or *H–2K–IA–IB* plus *H–2D* subregions of the *H–2^b* but not *H–2^d* nor *H–2^k* haplotypes are recognized by these anti-self H–2 antisera. The lack of reactivity with cells of the *H–2^d* or *H–2^k* haplotypes excludes many public specificities of *H–2^b* haplotype from these antisera. We have also tested spleen cells of SWR (*H–2^q*), RIII/2J (*H–2^r*), PL (*H–2^u*) and B10.M (*H–2^f*) mice for sensitivity to lysis by these antisera, because these haplotypes also share public specificities[17] with *H–2^b*. As none of these spleen cells was lysed to an appreciable extent by anti-self H–2 antisera, we conclude that the major reactivity of these antisera is directed to the private specificities controlled by *H–2^b*.

Molecular identity of species recognized by anti-self H–2 antisera

To identify the molecule(s) on spleen cells that was recognized by these antisera, cells from B10, B10.A(2R), B10.A(5R) and B10.D2n were radiolabelled by the ^{125}I-lactoperoxidase technique, lysates prepared and immunoprecipitated with anti-self H–2 antiserum or normal mouse serum. Immunoprecipitates were electrophoresed through polyacrylamide gels containing SDS, and the results analysed autoradiographically. Each of the antisera showed strong reactivity with molecules of molecular weights (MW) ~49,000 and ≤13,000 (the dye front) on B10 or B10.A(5R) cells (Fig. 4). One antiserum, which reacted well in cytotoxic tests with B10.A(2R) cells, showed strong precipitating activity for molecules of MW 49,000 and ≤13,000 on these cells, whereas another antiserum (S64) that had a lower cytotoxic titre (1 : 40) on B10.A(2R) cells precipitated less material of these molecular weights. Reactivity for molecules of MW 37,000 was also noted with B10 and B10.A(5R) cells. No reactivity was observed with any molecule on B10.D2n cells.

Because the major molecules precipitated by these antisera co-migrate with the H–2K or H–2D heavy chains[19], we sought to confirm that the major protein recognized by anti-self H–2 sera was the *H–2K*^{b} product. We did this by removing the H–2K^b molecule from lysates by immunoprecipitation with conventional antisera, and then determining if molecules recognized by anti-self H–2 sera remained in the lysates. The results of this 'preclearing' experiment, presented in Fig. 5, demonstrate that preclearing of B6 radiolabelled lysates with conventional anti-H–2K–IA–IB^b antisera removed the major molecules reactive with anti-self H–2 sera, whereas preclearing with normal mouse serum (Fig. 5) or conventional anti-H–2^k (data not shown) did not. Therefore we conclude that the major species recognized by the antiserum of mouse S64 [(C3H×B6)F$_1$ anti-B6T1] on normal B6 spleen cells is the molecule carrying the H–2K^b transplantation antigens. Similar experiments have confirmed the identity of the principal molecules recognized by other anti-self H–2 antisera as being the *H–2K*^{b} product.

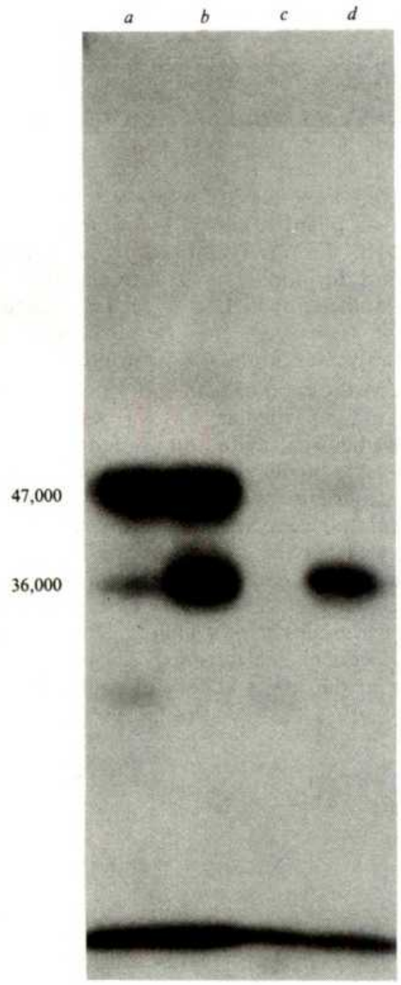


Fig. 5 Hybrid anti-self H–2 antisera and conventional anti-H–2K^b antisera recognize the same molecules. A single-cell suspension of B6 spleen cells was prepared and iodinated by the lactoperoxidase method. The cells were lysed and precleared with 3 ml of a 10% suspension of *S. aureus* bacteria as described in Fig. 4 legend. The lysate was then divided into two equal parts. One aliquot (lanes *a*, *b*) was incubated with 25 µl of normal mouse serum and the other (lanes *c*, *d*) was incubated with 25 µl of conventional anti-H–2K^b serum [(A×B10.D2)F$_1$ anti-B10.A(5R)]. The immune complexes were removed by the addition of 600 µl of the bacterial suspension followed by centrifugation. The supernatant of each reaction was further subdivided into two equal portions and incubated with 25 µl of anti-H–2K^b antiserum (lanes *a*, *c*) or 25 µl of (C3H×B6)F$_1$ antiserum (lanes *b*, *d*). These immunoprecipitation reactions were analysed by SDS-polyacrylamide gel electrophoresis as described in Fig. 4 legend.

Discussion

We have demonstrated that hybrid mice can generate a brisk humoral immune response to the major histocompatibility antigens presented on a parental virally induced leukaemia, although the hybrid mice also express those same histocompatibility antigens on their somatic cells. These results establish unambiguously the existence of autoreactive anti-MHC B-cell clones in hybrid mice. Earlier studies showed that not all immunizations of hybrid mice with parental tumours elicit anti-self H–2 antibodies. The Moloney cell-surface antigen (MCSA)[20], the X.1 antigen[21], and the Meth A antigen[22] were each detected by antisera from F$_1$ mice immunized with parental tumour cells, and no anti-self H–2 antibodies were detected in those sera. Anomalous anti-H–2 reactions are not without precedent, however, as demonstrated by the recent description of anti-H–2^k cytotoxic antibodies in the sera of mutant BALB/c mice immunized with normal BALB/c lymphoid cells[23].

The phenomenon of hybrid humoral response to parental antigens is reminiscent of the phenomenon of hybrid resistance to parental tumour grafts[10]. Two major features distinguish the humoral response described here from the *in vivo* phenomenon of F$_1$ rejection of parental bone marrow grafts or tumour grafts. First, the antigens detected on B6 cells by hybrid resistance map to the *H–2D* region, whereas those detected by anti-self H–2 antibodies map to the *H–2K* and *H–2D* regions. Second, the antigens detected in this serological study are expressed on heterozygous cells, whereas those detected in graft rejection do not seem to be. Furthermore, we did not detect anti-H–2^b antibody in (C3H×B6)F$_1$ hybrid mice immunized with B6 bone marrow, a strain combination that displays hybrid resistance[11].

Recent work on the induction of F$_1$ anti-parent cytotoxic T-lymphocyte responses by co-cultivation of F$_1$ spleen cells with homozygous parental cells[24,25], by treatment with allogeneic effect factor[26], or spontaneously[24,25] provide considerable evidence for the existence of autoreactive cytotoxic T-cells in the generation of cellular immune responses. Although the antigens recognized by F$_1$ anti-parent cytotoxic T-lymphocytes mapped to *H–2K* or *H–2D*, this apparent autoreactivity may reflect an H–2 compatibility requirement in conjunction with some other, perhaps endogenous viral, antigen and not cytotoxic function directed to the MHC product itself. Our present serological observations have no such ambiguity, and clearly demonstrate the possibility of humoral immune functions directed to self MHC products.

Ample evidence supports the physical and functional association of *H–2* products with oncornavirus glycoproteins[27–30], and it may be that the H–2 molecules of the tumour are recognized as foreign only in conjunction with unusual viral determinants. The failure to obtain anti-self H–2 antibodies from B6 mice immunized with this tumour might reflect the need for a proper immune response gene to recognize the unusual viral determinant. Note, however, that B6 mice are quite capable of responding to several viral antigens expressed on these cells[15]. Moreover, when anti-self responses were observed in hybrid mice, such anti-self H–2 antibodies were the major cytotoxic activities of the serum as demonstrated by absorption tests.

The observation here that normal bone marrow immunization did not elicit anti-self H–2 antibodies and the failure of earlier workers to observe such antibodies[20,21] in F$_1$ antiviral tumour immunizations may indicate that the anti-self response is only triggered by some novel, non-viral property of this tumour cell. Considered in that context, the failure of homozygous B6 mice to produce anti-self H–2 antibodies raises two questions regarding this response. First, whose antibodies are reactive with the H–2^b molecules, those of C3H or B6? Second, which genes must be heterozygous for this response to occur? It is possible, although unlikely, that self-reactive antibodies are absent from the repertoire of immunoglobulin genes of an inbred strain. More intriguing however, is the possibility that autoreactivity is itself generated by MHC heterozygosity. The resolution of the questions posed by the observation of anti-self H–2 antibodies will be of importance in the study of the generation and regulation of the immune response.

We thank P. Jelen and J. Timmins for assistance. R.R. is the recipient of a Leukaemia Society of America scholarship. This work was supported by grants CA-07175 and C1-22443 from the NCI.

Received 7 July; accepted 2 December 1980.

1. Snell, G. D. *J. Genet.* **49,** 87–105 (1948).
2. Gorer, P. A. *J. Path. Bact.* **44,** 691–697 (1937).
3. Medawar, P. *J. Anat.* **78,** 170–199 (1944).
4. Mitchison, N. A. *J. Genet.* **51,** 406–420 (1953).
5. Burnet, F. M. *The Clonal Selection Theory of Acquired Immunity* (Cambridge University Press, London, 1959).
6. Gorer, P. A. *Br. J. exp. Path.* **18,** 31–36 (1937).
7. Snell, G. D. *J. natn. Cancer Inst.* **20,** 787–824 (1958).
8. Hëllstrom, K. E. *Science* **143,** 477–478 (1964).
9. Boyse, E. A. *Immunology* **2,** 170–181 (1959).
10. Cudkowicz, G. & Stimpfling, J. H. *Immunology* **7,** 291–305 (1964).
11. Cudkowicz, G. in *The Proliferation and Spread of Neoplastic Cells 21st A.M.D. Anderson Symp. Cancer Res.* 661–641 (Williams & Wilkins, Baltimore, 1968).
12. Shearer, G. M. & Cudkowicz, G. *Science* **190,** 890–892 (1975).
13. Ishikawa, H. & Dutton, R. W. *J. Immun.* **122,** 529–536 (1979).
14. Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplantn Rev.* **29,** 89–124 (1976).
15. Risser, R., Stockert, E. & Old, L. J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3918–3922 (1978).
16. Bailey, D. W. *Transplantation* **11,** 325 (1971).
17. Altman, P. L. & Katz, D. D. (eds) *Inbred and Genetically Defined Strains of Laboratory Animals*, 122–140 (Fedn. Am. Soc. exp. Biol., Bethesda, 1979).
18. Old, L. J., Boyse, E. & Stockert, E. *Nature* **201,** 777–779 (1964).
19. Schwartz, B. D., Kato, K., Cullen, S. E. & Nathenson, S. G. *Biochemistry* **12,** 2157–2164 (1973).
20. Klein, E. & Klein, G. *J. natn. Cancer Inst.* **32,** 547–568 (1964).
21. Takahashi, I., Old, L. J. & Boyse, E. A. *J. exp. Med.* **131,** 1325–1341 (1970).
22. DeLeo, A. B., Shiku, H., Takahashi, T., John, M. & Old, L. J. *J. exp. Med.* **146,** 720–734 (1977).
23. Ivanyi, P., van Mourik, P., Breuning, M. & Melief, C. *Immunogenetics* **10,** 319–332 (1980).
24. Cudkowicz, G., Nakano, K. & Nakamura, I. in *Immunopathology* (eds Milgrom, F. & Albini, B.) 31–36 (Karger, Basel, 1979).
25. Ishikawa, H. & Dutton, R. W. *J. Immun.* **125,** 656–662 (1980).
26. Altman, A. & Katz, D. H. *Immun. Rev.* **51,** 3–34 (1980).
27. Chesebro, B., Wehrly, K. & Stimpfling, J. *J. exp. Med.* **140,** 1457–1467 (1974).
28. Blank, K. J., Freedman, H. A. & Lilly, F. *Nature* **260,** 250–251 (1976).
29. Zarling, D., Keshet, I., Watson, A. & Bach, F. *Scand. J. Immun.* **8,** 497–508 (1978).
30. Bubbers, J. E. & Lilly, F. *Nature* **269,** 808–809 (1977).
31. Marchalonis, J. J. *Biochem. J.* **113,** 299–305 (1969).
32. Phillips, D. R. & Morrison, M. *Biochem. biophys. Res. Commun.* **40,** 284–289 (1970).
33. Van Zaane, D., Gielkens, A. L. J., Dekken-Michielsen, M. J. A. & Bloemers, A. P. J. *Virology* **67,** 544–552 (1975).
34. Kessler, S. W. *J. Immun.* **115,** 1617–1624 (1975).
35. Brugge, J. S. & Erikson, R. L. *Nature* **269,** 346–348 (1977).
36. Laemmli, U. K. *Nature* **227,** 680–685 (1970).

LETTERS

γ-Ray observations of Cygnus X-3 at energies of 10^{12}eV

S. Danaher, D. J. Fegan & N. A. Porter

Physics Department, University College, Dublin, Ireland

T. C. Weekes

Mount Hopkins Observatory, Harvard-Smithsonian Center for Astrophysics, Amado, Arizona 85640

The Crimean Observatory group have detected γ rays from the direction of the X-ray source Cyg X-3 (refs 1–5): rays of energy 10^{12}eV were observed using the atmospheric Cerenkov technique. Measurements with satellite-borne instruments in the 10^8 eV region have shown effects from Cyg X-3 in one case[6], but not in another[7]. Similarly, γ rays have been seen in some balloon flights[8], but not in those of McKechnie *et al.*[9] nor of White *et al.*[10]. A study using the atmospheric Cerenkov technique at 10^{11} eV showed no effect[11], though it has been shown to be compatible with the Crimean result[4]. We report here observations taken at Mount Hopkins Observatory, Arizona (altitude 2.3 km) in April, May and June 1980 using a similar system to that of Stepanian *et al.*[3], which seem to confirm, at least qualitatively, the Crimean results.

The Crimean group found a periodic signal with the same 4.8-h period as the X rays, and having two peaks: a narrow main one centred at 0.18 periods after X-ray minimum, and a broader interpulse at 0.7–0.8 periods after X-ray minimum. They have also reported a sporadic non-periodic effect. Recently (Stepanian, personal communication) they have found that the interpulse peak dominates over the main one.

In the present series of observations two mirrors of diameter 1.5 m were used, with a 5-cm photomultiplier and aperture stop at the focus of each, giving a full field of view of 2.0°. Coincidences were taken between the two detectors with a resolving time of 7 ns, and were recorded on magnetic tape together with pulses from the single detectors, and random coincidences. Current controlled servo lamps kept the background light constant. Observations were made in drift scans lasting 30 min each: 10 min before the source entered the field (OFF), 10 min in the field (ON), and a final 10 min outside the field (OFF). A scan was rejected if (1) the OFF parts of the scan divided into 120 10-s intervals failed a Poisson homogeneity test at the 1% level, (2) the two OFF observations differed by more than two Poisson standard deviations, or (3) if in any 10-s interval, the count rate was sufficiently high to give a 4σ departure from the mean. Of 139 scans on Cyg X-3, 95 were acceptable. These were taken at various phases on the 4.8-h period.

Figure 1 shows the ratio of ON/OFF rates plotted in 10 bins corresponding to the particular phases of X-ray emission. Standard errors have been calculated experimentally, for each phase bin and are ~14% higher than the Poisson values, consistent with previous experience with the Cerenkov technique. Phases have been calculated using Stepanian's constants and formulae. The ratios are clearly close to unity except at the interval from 0.7 to 0.8, which has a 3.5σ deviation from unity. The number of counts involved are: ON 799, OFF 662 before and 668 after transit. Three drift scans were available. Random coincidences and single rates were normal during the scans. The apparent flux corresponding to the effect at the interpulse is $\sim 1.5 \times 10^{-10}$

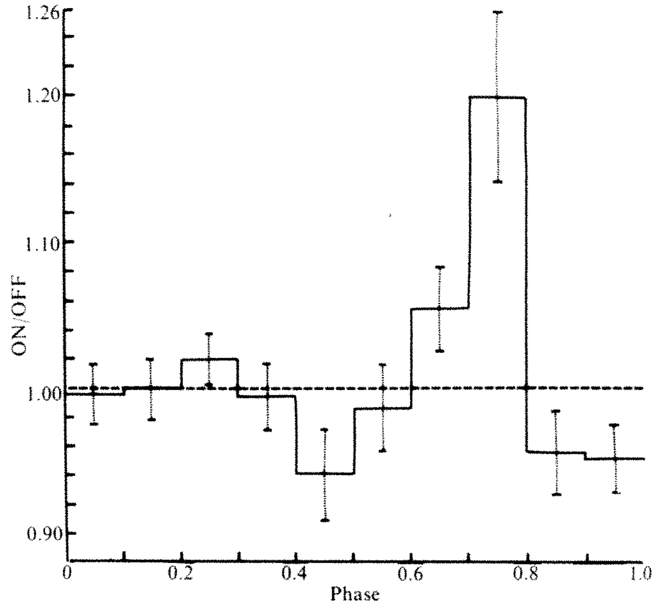


Fig. 1 Phase histogram of γ-ray emission from Cyg X-3.

photons cm^{-2} s^{-1} in good agreement with the value of 9×10^{-11} photons cm^{-2} s^{-1} quoted in ref. 4, and the threshold energy is $\sim 2\times10^{12}$eV, similar to that of the Crimean group.

While there is some slight uncertainty in the period and period derivative for Cyg X-3, (refs 12, 13), our phases are directly comparable with the Crimean results as we have used the same constants. We therefore seem to have confirmed, at least qualitatively, their observations. Note that the phase of the peak in our results corresponds to maximum emission for lower energy γ rays[6] and for X rays[14]. If the Cerenkow measurements are accepted as indicating a flux of high energy γ rays, they favour models which require a young pulsar[15]. A periodic analysis of the high-energy data may reveal the predicted 4–20 ms rotation.

We thank Dr A. A. Stepanian for communicating results before publication, the National Board for Science and Technology of Ireland, and the Smithsonian Scholarly Studies Fund, for partial support.

Received 20 October; accepted 15 December 1980.

1. Vladimirsky, B. M., Stepanian, A. A. & Fomin V. P. *Proc. 13th int. Conf. Cosmic Rays* **1**, 456–460 (1973).
2. Vladimirsky, B. M., Neshpor, Yu. I., Stepanian, A. A. & Fomin, V. P. *Proc. 14th int. Conf. Cosmic Rays* **1**, 118–122 (1975).
3. Stepanian, A. A., Vladimirsky, B. M., Neshpor, Yu. I. & Fomin, V. P. *Proc. 15th int. Conf. Cosmic Rays* **1**, 135–140 (1977).
4. Neshpor, Yu. I. *et al. Astrophys. Space Sci.* **61**, 349–355 (1979).
5. Mukanov, J. B., Nesterova, N. M., Stepanian, A. A. & Fomin, V. P. *Proc. 16th int. Conf. Cosmic Rays* **1**, 143–146 (1979).
6. Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thompson, D. J. *Astrophys. J. Lett.* **212**, L63–L66 (1977).
7. Bennett, K. *et al. Astr. Astrophys.* **59**, 273–274 (1977).
8. Galper, A. M. *et al. Proc. 15th int. Conf. Cosmic Rays* **1**, 131–134 (1977).
9. McKechnie, S. P., Mount, K. E. & Ramsden, D. *Astrophys. J. Lett.* **207**, L151–L154 (1976).
10. White, R. S. *et al. Nature* **284**, 608–610 (1980).
11. Weekes, T. C. & Helmken, H. F. *Astrophys. J.* **228**, 531–5 (1979).
12. Mason, K. O. & Sanford, P. W. *Mon. Not. R. astr. Soc.* **189**, 9P–14P (1979).
13. Elsner, R. F. *et al. Astrophys. J.* **239**, 335–344 (1980).
14. Bonnet-Bidaud, J. M., Ilovaisky, S., Manzo, G., Molteni, D. & Robba, N. R. *Astr. Astrophys.* **62**, 275–276 (1978).
15. Bignami, C. F., Maraschi, L. & Treves, A. *Astr. Astrophys.* **55**, 155–158 (1977).

Effect of a surrounding star cluster on a magnetoid

Martin Lugg

Department of Physics, University College of Wales, Penglais,
Aberystwyth SY23 3BZ, UK

Magnetoids—supermassive rotating highly magnetic stars—have long been considered[1–4] as models of the power sources of active galactic nuclei, but previous studies have taken no account of interaction with neighbouring matter. We now show that interaction with a star cluster (of a kind likely to exist near the nucleus of a galaxy) will disrupt the main body of a nuclear magnetoid forming a dense core which is liable to explosion and/or gravitational collapse. The lifetime of such a magnetoid is estimated at 10^4 yr, too short to explain the activity of quasars and Seyfert galaxies (which might, nevertheless, be accounted for by residual black holes).

A magnetoid is assumed to form from gas falling into the central potential well of a galaxy and collapsing until a quasi-equilibrium state is reached, with radiation pressure, centrifugal and magnetic forces balancing gravity. Subsequent evolution involves a slow contraction during which gravitational potential energy is converted to rotational energy, which is then lost through equatorial mass shedding and magnetic-dipole radiation. Typical magnetoid parameters required to explain highly active nuclei are a mass $M\sim10^8M_\odot$ and a radius $R\sim10^{16}$ cm, giving a mean density $\bar\rho_m\sim10^{-8}$ g cm^{-3}.

Previous investigations have considered the magnetoid mainly as an isolated body, with little attention being paid to its surroundings, in particular the dense star cluster contained in the centre of the parent galaxy. Our Galaxy, for example, has a stellar core of density $\sim10^7M_\odot\,\mathrm{pc}^{-3}$ out to $\gtrsim0.1$ pc (ref. 5), values which should at least be matched in more active nuclei. Such a star cluster will have a dramatic effect on a recently formed magnetoid.

Here we assume the star cluster has an isotropic velocity distribution so that the collision frequency of stars with the magnetoid is $f=n_*\sigma V_0$, where n_* is the number density of stars, V_0 is their velocity dispersion and $\sigma(=\pi r_0^2)$ is the collision cross-section; r_0 is the impact parameter at which an incoming star will just graze the magnetoid. We take as typical a cluster of size $r_c\sim1$ pc containing $n_*\sim10^7$ stars pc^{-3} of average mass $m_*\sim1M_\odot$ surrounding a magnetoid of mass $10^8M_\odot$; then $V_0\simeq(GM/r_c)^{1/2}\simeq6\times10^7$ cm s^{-1} and the collision frequency

$$f\simeq200\,M_8^{1/2}R_{16}\quad\mathrm{yr}^{-1}\tag{1}$$

with M_8 being the magnetoids mass in units of $10^8M_\odot$ and R_{16} its radius in units of 10^{16} cm. A colliding star on its passage through the magnetoid will experience a drag force of

$$F_*=m_*\dot V_*\simeq\rho_m\pi r_*^2v_*^2\tag{2}$$

r_*, v_* are the radius and velocity of the star respectively. We take the magnetoid to be modelled by a polytrope of index 3, as usual for supermassive stars, and use the density function $\rho_m(r)$ computed by Monaghan and Roxburgh[6]. This model gives a central condensation, $\rho_c/\bar\rho_m=54.2$, thinning to a large diffuse envelope with $\rho_m<\rho_c/100$ for $r>0.58R$. Equation (2) can be written in the form

$$\frac{\mathrm{d}E_*}{\mathrm{d}s}=\frac{-2\pi}{m_*}\rho_m(s,x)E_*r_*^2\tag{3}$$

where s is the distance a star has travelled in the magnetoid, and x is the impact parameter. Using a polynomial fit to $\rho_m(s,x)$ equation (3) was integrated to obtain the energy lost by a star in one passage of the magnetoid, ΔE_*, as a function of impact parameter. The percentage energy loss reached a maximum of 15% for a star passing through the centre of the magnetoid.

The equation for ΔE_* was then integrated over x, assuming that the stars collide evenly over the magnetoid cross-section. This gave the average percentage energy loss per collision as 0.96%. The energy loss $\Delta E_*(x)$ equals the average $\langle\Delta E_*\rangle$ at $x=0.43R$ so 18% of the stars lose more than 1% of their impact energy in one collision. With an impact energy $E_i\simeq GMm_*/R$ the average energy loss per collision is

$$\langle\Delta E_*\rangle=2.6\times10^{49}M_8^2R_{16}^{-3}(r_*/r_\odot)^2\quad\mathrm{erg}\tag{4}$$

spread over a crossing time of $\lesssim1$ yr. This energy is deposited in the magnetoid at a rate given by equations (1) and (4) as

$$L_*=f\langle\Delta E_*\rangle\simeq3.7\times10^{44}M_8^{5/2}R_{16}^{-2}(r_*/r_\odot)^2\quad\mathrm{erg\,s^{-1}}\tag{5}$$

which is small compared with the magnetoid's thermal luminosity, L_{th}. For a supermassive star radiation pressure dominates gas pressure and the thermal luminosity is given by the Eddington critical luminosity, L_{Edd}, at which radiation pressure and gravity balance,

$$L_{\mathrm{th}}=L_{\mathrm{Edd}}=1.3\times10^{46}M_8\quad\mathrm{erg\,s^{-1}}\tag{6}$$

Since $\langle\Delta E_*\rangle\simeq10^{-2}E_i$ a star requires an average of ~100 collisions before its orbit lies entirely within the magnetoid. However, $\langle\Delta E_*\rangle$ is much greater than the star's original kinetic energy in the cluster, $E_0=\frac{1}{2}m_*V_0^2\simeq2.5\times10^{48}$ erg, so that after one collision a star becomes bound to the magnetoid and collides on every subsequent orbit to become captured after ~100 orbits or $\sim10^4$ yr.

The initial capture rate of $\gtrsim100$ stars yr^{-1} can only be maintained while there are stars in low angular momentum orbits; when these orbits become depleted the subsequent capture rate depends on the rate of refilling them by angular momentum diffusion in the cluster. The orbits get refilled in a

0028-0836/81/070569-02$01.00

time not much shorter than the relaxation time scale, $t_r \simeq V_0^3/G^2 m_*^2 n_* \log(N/2) \simeq 10^9$ yr (ref. 7), required to randomize the orbits completely. This is much greater than the 10^4 yr required to capture a star, so once the low angular momentum orbits have been emptied the capture rate drops sharply. The collision cross-section of the magnetoid, $\sigma = \pi r_0^2 = 2\pi GMR/V_0^2 \simeq 3.5 \times 10^{-2}\,\mathrm{pc}^2$, so for an isotropic cluster of radius ~1 pc the magnetoid will capture an estimated fraction 1/100 of the cluster, or $\sim 10^5$ stars, before the capture rate drops.

We now consider the effect of these captured stars on the magnetoid as they continue to lose energy and sink towards the magnetoid's centre. Ablation of a stellar envelope occurs when $\rho_m(r) V_*^2 > \rho_e V_{esc}$ (ref. 9), where ρ_e is density of the stellar envelope and V_{esc} the escape velocity from the star, that is when

$$\rho_e < 1.7 \times 10^{-3} (V_*/V_i)^2 (\rho_m(r)/\rho_c)(M_*/M_\odot)^{-1}(r_*/r_\odot) \quad \mathrm{g\,cm^{-3}} \tag{7}$$

where V_i is the impact velocity of a star with the magnetoid, about $10^9\,\mathrm{cm\,s^{-1}}$. A solar-type star passing through the centre of the magnetoid, thus loses the part of its envelope rarer than $1.7 \times 10^{-3}\,\mathrm{g\,cm^{-3}}$, which is the outer 0.1 $r_\odot$, or $10^{-4} M_\odot$. A red giant[8] of $1.3 M_\odot$, however, would be stripped down to around one-sixth of its original radius and one-third of its original mass. The detailed picture will depend on how evolved the galactic nucleus is at the time of magnetoid formation, but in any case dwarf and sub-dwarf stars will outnumber giants, and ablation is not likely to be important except near the centre of the magnetoid. In a counter-effect, as the dwarf stars slow down they will accrete mass from the magnetoid envelope at a rate[9]

$$A \sim \frac{2\pi (G m_*)^2 \rho_m}{V_*^3} \sim 10^{-5} (m_*/m_\odot)^2 (V_*/10^8\,\mathrm{cm\,s^{-1}})^{-3} M_\odot \quad \mathrm{yr^{-1}} \tag{8}$$

and carry it further into the magnetoid, though this is negligible for $V_* \gtrsim 10^7\,\mathrm{cm\,s^{-1}}$.

Another possible effect will be fragmentation of the gas in the magnetoid due to the differential gravitational pull of the stars. This can occur when the average density in the form of stars, $\bar{\rho}_*$, exceeds the gas density, $\bar{\rho}_m$, in a certain region[10]; that is, when 10^5 solar-type stars are confined within a radius r_c from the magnetoid's centre where $r_c = 2.6 \times 10^{14} M_8^{-1/3} R_{16}$ cm. To reach this state a star has to lose 96% of the energy it had when initially captured, a process taking only $\sim 10^3$ yr.

Finally the stars will be colliding with each other at a rate $f_c \simeq 2\pi N n_* r_*^2 V$. For 10^5 stars confined within a radius r_c, $V \simeq 10^8 (r_c/10^{15}\,\mathrm{cm})^{-1/2}\,\mathrm{cm\,s^{-1}}$, so $f_c \simeq 200\ (r_c/10^{15}\,\mathrm{cm})^{-7/2}\,\mathrm{yr^{-1}}$. The energy released ($\sim 10^{48}$ erg per collision) will tend to disrupt the magnetoid and aid fragmentation, while the debris will help to slow down other incoming stars.

The net effect of these processes is to concentrate mass, in the form of a dense core of coalescing stars and gas, in the centre of the magnetoid. For a star cluster initially having an isotropic velocity distribution this core will have insufficient angular momentum to prevent it collapsing to normal stellar densities and beyond. Collapse may be stopped by a nuclear explosion if the core has a mass $\lesssim 1.5 \times 10^5 M_\odot$ (ref. 11), otherwise the core collapses to a black hole.

It thus seems inevitable that a magnetoid (or any other form of supermassive star) formed in an ordinary galactic nucleus will be disrupted in a time $\sim 10^4$ yr by the processes described, to leave a dense star cluster (the higher angular momentum part of the original cluster) surrounding a black hole of either $\sim 10^5$–$10^6 M_\odot$ if no core explosion occurred, or $\sim 10^2$–$10^3 M_\odot$ as a remnant of an explosion. Such a black hole will then grow and be able to power an active galactic nucleus[12] for the $> 10^6$ yr required for Seyfert galaxies and quasars.

I thank Drs J. Frank and D. F. Falla for helpful discussions, and acknowledge receipt of a Research Studentship from the SRC.

Received 17 June; accepted 2 December 1980.

1. Ozernoy, L. M. & Usov, V. V. *Astrophys. Space Sci.* **13,** 3–35 (1971).
2. Ozernoy, L. M. & Usov, V. V. *Astrophys. Space Sci.* **25,** 149–194 (1973).
3. Ozernoy, L. M. & Usov, V. V. *Astr. Astrophys.* **56,** 163–172 (1977).
4. Kundt, W. *Astrophys. Space Sci.* **62,** 335–345 (1979).
5. Bailey, M. *Mon. Not. R. astr. Soc.* **190,** 217–225 (1980).
6. Monaghan, J. J. & Roxburgh, I. W. *Mon. Not. R. astr. Soc.* **131,** 13–22 (1965).
7. Spitzer, L. & Harm, R. *Astrophys. J.* **127,** 544–553 (1958).
8. Allen, C. W. in *Astrophysical Quantities,* 205 (Athlone, London, 1963).
9. Bondi, H. & Hoyle, F. *Mon. Not. R. astr. Soc.* **104,** 273–284 (1944).
10. Safronov, V. *Soviet Phys. Doklady* **5,** 13–16 (1960).
11. Zel'dovich, Ya. B. & Novikov, I. D. in *Relativistic Astrophysics* Vol. 1, 382 (University of Chicago Press, 1971).
12. Hills, J. G. *Nature* **254,** 295–298 (1975).

Atmospheric transport of soil dust from Africa to South America

J. M. Prospero, R. A. Glaccum & R. T. Nees

Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149

The arid and desert regions of North Africa are a prolific source of atmospheric dust. This dust is, for example, responsible for the 'red snows' reported in the Alps and Pyrenees[1] and for dust falls further north in Europe[2–6], but these phenomena are infrequent and sporadic. By contrast, the transport of mineral dust into the tropical North Atlantic is common and often produces a widespread dense haze[7]. Investigations in the past 15 yr[8–11] have shown that, in the summer months[12,13], the distribution of dust is related to macro and micro-meteorological circumstances, that the dust often reaches altitudes of 5–7 km and that it may be spread over several hundreds of kilometres in latitude and extend to the Caribbean Sea and the south-east United States. We report here the results from an aerosol sampling station at Cayenne, French Guiana, which indicate that large quantities of soil dust are being carried out of North Africa and across the Atlantic during the winter months as well but at this time of year the transport is primarily to South America.

A seasonal meteorological control on African dust transport can be inferred from a 15-yr mineral aerosol study[10] carried out on Barbados, West Indies (13°10′N, 59°30′W). There is a pronounced annual cycle in dust concentration with the maximum occurring during the summer; monthly mean winter concentrations are 10–100 times lower. The seasonal cycle in dust concentration at Barbados is apparently not due to any cessation of dust transport out of Africa during the winter months—the dust haze along the west coast of Africa occurs more frequently in winter than in the summer[14]. However, the winter haze distribution is shifted about 10° to the south of the

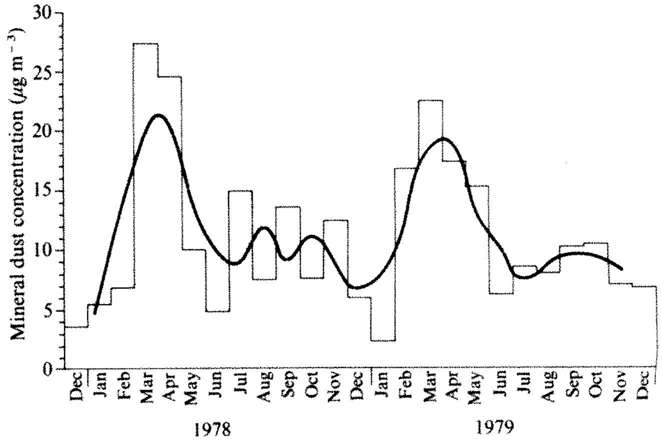


Fig. 1 Monthly arithmetic mean mineral dust concentration in the surface-level air at Cayenne, French Guiana. The heavy line is the three month moving average. The means are computed from nominal 24-h daily samples.

Table 1 Mineral composition of soil aerosols

a Cayenne aerosols: average composition of grouped samples selected from a set of 29					Mineral composition*						
	Mica	Kaol	Chl	Qtz	Mic	Plg	Cal	Gyp	Gib	Gth	No. of samples
Assumed origin:											
North Africa	62	10	5	8	2	5	1	2	1	4	13
South America	26	5	4	21	1	8	5	1	15	14	12
b Samples collected in the aerosol network during the dust outbreak of 20–26 March 1978											
Dakar 21–22 March	57	7	7	14	2	5	5	1	0	2	
Barbados 25–27 March	60	6	4	10	3	6	5	3	0	2	
Cayenne 25–26 March	59	8	5	10	3	5	4	3	1	3	

* Mica, mica and illite; Kaol, kaolinite; Chl, chlorite; Qtz, quartz; Mic, microcline; Plg, plagioclase; Cal, calcite; Gyp, gypsum; Gib, gibbsite; Gth, goethite.

summer distribution. Thus, this seasonal cycle at Barbados could be due to the seasonal southward shift of the large-scale circulation patterns over Africa and the tropical Atlantic. The continued dust activity in Africa during the winter and the consequent haze distribution suggested that the principal dust transport trajectories might be found in the latitudes south of that of Barbados. To test this hypothesis, we established an aerosol sampling station in Cayenne, French Guiana (4°50'N, 52°22'W) in December 1977; the station has been in continuous operation ever since then.

Monthly mean mineral aerosol concentrations at Guiana (Fig. 1) were determined from daily filter samples collected at a coastal site by weighing the residue after extracting the filter with water and ashing at 500 °C (ref. 9). The greatest mineral aerosol concentrations are observed early in the year with the maximum monthly mean occurring in March in both 1978 and 1979. The Cayenne annual maximum values (29 $\mu g\ m^{-3}$ and 23 $\mu g\ m^{-3}$ for 1978 and 1979 respectively) are markedly higher than the annual maximum monthly means for Barbados which were in the range 15–18 $\mu g\ m^{-3}$ between 1976 and 1979; the annual maximum at Barbados occurred during June, July or August.

The day-to-day variability of the mineral dust concentration in Cayenne is quite large. This can be seen in Fig. 2 which presents data for March 1978. Similarly large variations have been measured at other stations in the tropical North Atlantic during the summer[9,10,15,16].

We analysed the mineralogical composition of 29 aerosol samples by X-ray diffraction[11]. The samples were selected to represent the periods of highest dust concentration each month. The samples can be divided into three mineral groups. One group of 13 samples has the same composition as the dust from North Africa previously collected from ships and island stations in the tropical North Atlantic[11]. The second group of 12 samples has a distinctly different composition from that of African dust and which we attribute to the presence of soil material derived from the land mass of South America, most likely Brazil. The mean composition of these two groups is presented in Table 1*a*. The third group of four samples has a composition that suggests that the samples are a mixture of typical North African and South American soil material.

Using satellite imagery we focused on the two dust concentration maxima occurring on 21–22 and 24–25 March (Fig. 2). A search of GOES satellite imagery at the National Environmental Satellite Service yielded a sequence of photographs showing two dust outbreaks emerging from the coast of Africa, a relatively weak one on 16–17 March and a major one on 20–22 March; the dust clouds could be followed across the Atlantic in the photographic sequence. The transit time across the Atlantic is 4–5 days, and the distance covered ~4,000 km.

The passage of the dust clouds was detected at several stations in an aerosol network established in the tropical North Atlantic since 1974. At Dakar, Senegal (14°44'N, 17°33'W), the dust concentration reached a maximum of 580 $\mu g\ m^{-3}$ on 21 March, the day on which the satellite photographs show the main plume emerging from the coast. The outbreak of 17 March yielded a dust load of 270 $\mu g\ m^{-3}$ at Dakar.

Our long-term studies at Barbados indicate that very high concentrations would not be expected at this time of year. In fact, the first of the two dust outbreaks passed to the south of the island, and there was no large increase in dust load on 21 and 22 March. However, on 25 and 26 March the dust load increased dramatically from 6 to 104 $\mu g\ m^{-3}$; on 27 March it dropped to 5 $\mu g\ m^{-3}$. Focusing on the second and largest outbreak, we selected from each station the filter collected during the peak dust concentration period and analysed it by X-ray diffraction (Table 1*b*). Clearly, the dust mineralogy at Cayenne is essentially the same as that of the samples from the other stations. The similarities both in the mineralogies of the Cayenne and Barbados dust samples in this one dust episode, and between this one Cayenne sample and the bulk of those collected during the winter and spring in Cayenne suggest that these aerosols are indeed of African origin.

The wind climatology of French Guiana is consistent with the observation of Saharan dust in winter and Brazilian and local dusts in summer. The large-scale circulation is principally determined by the position of the equatorial trough (that is, a latitudinal belt of relatively low pressure) commonly referred to as the intertropical convergence zone (ITCZ) which can be pictured as the area of convergence between the north-east and south-east trade wind regimes[17]. The ITCZ translates meridionally on a seasonal basis about its annual mean position at 5°N. In the winter months, when the ITCZ is in its most southerly position, Cayenne lies on the northern edge of the ITCZ causing the winds at Cayenne to be predominantly north-east (57%) and east (32%); they are also relatively steady with only a 0.4% incidence of calms[18]. In the summer, the ITCZ is in its most northerly position and Cayenne lies well within the ITCZ circulation; although winds are predominantly north-east

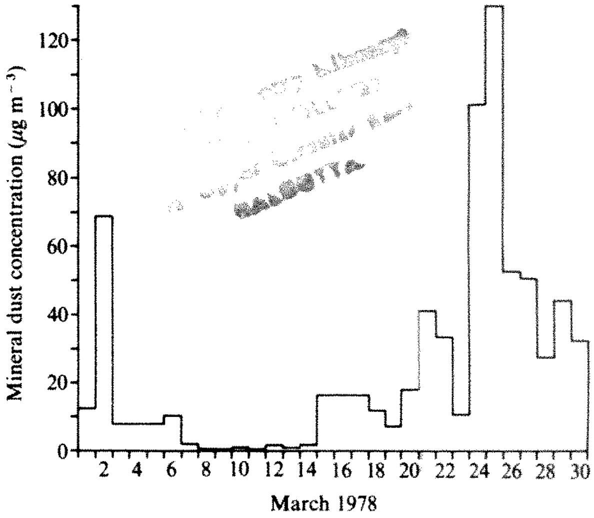


Fig. 2 Daily mineral dust concentration in surface level air at Cayenne, French Guiana, during March, 1978. The periods of relatively high dust concentration are all attributable to incursions of air parcels containing dust advected from North Africa.

(24%), east (41%) and south-east (18%), they are not so steady in speed or direction and there are significantly more periods of light and variable winds or of calm[18]. Saharan dust is occasionally transported into the Cayenne region during the summer on north-east winds, but the concentrations are never as great as during the winter. Of our 29 analysed samples, three of the four samples of mixed African and South American mineralogies were collected during the summer months.

The mineralogy of the aerosol samples placed in the South American group in Table 1*a* is consistent with the known characteristics of the soils in this region. The soils of French Guiana and much of the north-east are predominantly laterites formed by the intensive chemical weathering processes that are common to hot, humid, tropical climates. Such highly leached soils are typically characterized by the presence of high concentrations of iron-oxide minerals (geothite, haematite) and aluminum oxide minerals (gibbsite, boehmite), the relative amounts of which are determined by the chemical and mineral composition of the underlying bedrock. Indeed, the concentration of aluminum oxide minerals is often so great that gibbsitic bauxites are mined in French Guiana, Venezuela, Brazil and Surinam.

Much of the summer and autumn dust collected at Cayenne is relatively fine grained which suggests that it has been transported over a considerable distance. Given the dominance of the south-east trade winds in these seasons, this soil material is most likely derived from the eastern portions of Brazil. In the summer and autumn, the aerosol filters are often blackened by a component that seems to be elemental carbon from combustion sources. This material could come from local anthropogenic sources such as cars in the Cayenne area during periods of south and south-west winds; however, the population of Cayenne is rather small (the entire population of French Guiana was only 55,000 in 1975). We suspect that the black material might be a product of cut-and-burn agriculture which is common in the northeastern portion of South America, especially Brazil. Burning is most intensive in the dry season (late summer to autumn). The summer–autumn samples also often contain much organic material, especially plant fibre and insects. In contrast, the winter (African) dust is always very fine textured, has a beige colour characteristic of Saharan aerosols, and is essentially devoid of readily identifiable biological material.

Previous studies suggest that the African dust collected during the peak concentration period comes from the western and central Sahara and bordering regions to the south[19–25]. Although we have not attempted a detailed study of possible sources for the winter dust, the GOES satellite imagery for March 1978 suggests that these same areas might be active in the winter. There is also considerable dust storm activity in Chad and Niger during December to February[19–21,24,25]; as a consequence of these storms, extremely dusty air masses, known as the harmattan, penetrate to the Gulf of Guinea. The coincidence of maximum dust activity along the coast of the Gulf with the time of minimum dust transport to the region of French Guiana suggests that this massive transport of dust must be terminated by fairly complete removal, probably in the regions of high precipitation that are characteristic of the ITCZ.

The transport of dust out of Africa into and across the Atlantic has important implications. First, the quantity of dust being transported has been estimated to be in the range of 100–400 megatons (where 1 megaton = 10^{12} g) per yr (refs 13, 15, 26). Our results support the upper range of estimates. Such quantities of material could provide a very substantial fraction of the non-biogenic material entering the deep-sea sediments in the tropical and equatorial North Atlantic[12,15]. Second, 8.5% of the dust mass collected over the western Atlantic consists of particles (density, 2.5 g cm^{-3}) with diameters greater than 12 μm (ref. 27); the equivalent aerodynamic diameter of particles having a density of 1 g cm^{-3} would be 20 μm. Therefore many types of fungal spores (which usually have diameters between 5–20 μm) could be transported across the Atlantic in a viable state; such a transport might be of aerobiological significance. Finally, concentrations[26] of Saharan dust comparable with those measured in French Guiana can produce a marked increase in heating rates in the middle troposphere and a cooling at the surface[28]. This increases atmospheric stability to the extent that there may be a significant impact on the meteorological processes occurring over a large area of the topical and equatorial North Atlantic.

We thank R. Causse and C. Dardeau for supervising the sampling programme in Cayenne and D. Moreau and J. Rine for technical assistance. This work was supported by the Division of Atmospheric Sciences, NSF grants ATM76-02198 and ATM78-12246.

Received 2 September; accepted 20 October 1980.

1. Bücher, A. & Lucas, C. *La Meteorologie* **5**, 53–69 (1975).
2. Free, E. E. *The Movement of Soil Material by the Wind* (US Department of Agriculture, Washington DC, 1911).
3. *Nature* **291**, 112 (1968).
4. Pitty, A. F. *Nature* **220**, 364–365 (1968).
5. Stevenson, C. M. *Weather* **24**, 126–132 (1969).
6. Tullett, M. T. *Weather* **33**, 48–52 (1978).
7. Darwin, C. *Q. Jl Geol. Soc. Lond.* **2**, 26–30 (1846).
8. Delany, A. C. *et al. Geochim. cosmochim. Acta* **31**, 885–909 (1967).
9. Savoie, D. L. & Prospero, J. M. *J. geophys. Res.* **82**, 5954–5963 (1977).
10. Prospero, J. M. & Nees, R. T. *Science* **196**, 1196–1198 (1977).
11. Glaccum, R. A. & Prospero, J. M. *Mar. Geol.* **37**, 295–321 (1980).
12. Jaenicke, R. & Schutz, L. *J. geophys. Res.* **83**, 3585–3599 (1978).
13. Carlson, T. N. & Prospero, J. M. *J. appl. Met.* **11**, 283–297 (1972).
14. McDonald, W. F. *Atlas of Climatic Charts of the Oceans* (Department of Agriculture Weather Bureau, Washington DC, 1938).
15. Prospero, J. M. & Carlson, T. N. *J. geophys. Res.* **77**, 5255–5265 (1972).
16. Prospero, J. M., Savoie, D. L., Carlson, T. N. & Nees, R. T. in *Saharan Dust: Mobilization, Transport, Deposition* SCOPE rep. 14 (ed. Morales, C.) 171–186 (Wiley, New York, 1979).
17. Snow, J. W. in *Climates of Central and South America: World Survey of Climatology*, Vol. 12 (ed. Schwerdtfeger, W.) 295 (Elsevier, Amsterdam, 1976).
18. *Summary of Synoptic Meteorological Observations (SSMO), South America and Selected Island Coastal Marine Areas* Vol. 1, 80–158 (National Space Technical Laboratory, Mississippi, 1977).
19. Hamilton, R. A. & Archbold, J. W. *Q. Jl R. met. Soc.* **71**, 231–265 (1945).
20. Dubief, J. in *Saharan Dust: Mobilization, Transport, Deposition*, SCOPE Rep. 14 (ed. Morales, C.) 27–48 (Wiley, New York, 1977).
21. Kalu, A. E. in *Saharan Dust: Mobilization, Transport, Deposition*, SCOPE Rep. 14 (ed. Morales, C.) 95–118 (Wiley, New York, 1977).
22. Morales, C. in *Saharan Dust: Mobilization, Transport, Deposition*, SCOPE Rep. 14 (ed. Morales, C.) 119–131 (Wiley, New York, 1977).
23. Yaalon, D. H. & Ganor, E. in *Saharan Dust: Mobilization, Transport, Deposition*, SCOPE Rep. 14 (ed. Morales, C.) 187–193 (Wiley, New York, 1977).
24. Bertrand, J., Baudet, J. & Drochon, A. *J. Rech. atmos.* **8**, 845–860 (1974).
25. Domergue, J. L. thesis, Univ. Paul Sabatier de Toulouse (1980).
26. Carlson, T. N. & Benjamin, S. G. *J. atmos. Sci.* **37**, 193–213 (1980).
27. Savoie, D. thesis, Univ. Miami, (1978).
28. Carlson, T. N. *Mon. Weath. Rev.* **107**, 322–335 (1979).

Coloured materials and photoluminescence centres in anodic film on aluminium

Y. Yamamoto, N. Baba & S. Tajima

Laboratory of Electrochemistry and Inorganic Chemistry,
Tokyo City University, Tokyo 158, Japan

Anodic oxide films formed on aluminium in lower aliphatic carboxylic acids exhibit self-colouring of yellow through brown, and characteristic properties such as photoluminescence (PL), electroluminescence (EL), and electron spin resonance (ESR). These phenomena are closely associated with the carboxylate species incorporated in the oxide during anodization[1]. Although the presence of incorporated carboxylate species in some cases[2,3] and their distribution in oxalic acid film have been established[4], there have been little data available on the state of incorporated impurities in the oxide. We have studied the distributions of coloured material and PL centres in the oxide and report here that most of the coloured material is located in the outer region of the cell walls of an oxalic acid film, but that the PL centres are in the middle region and have different properties from those of the outer region. These results suggest that the incorporated impurities which are present in the various states in oxide have to be converted during anodization.

0028–0836/81/070572–03$01.00

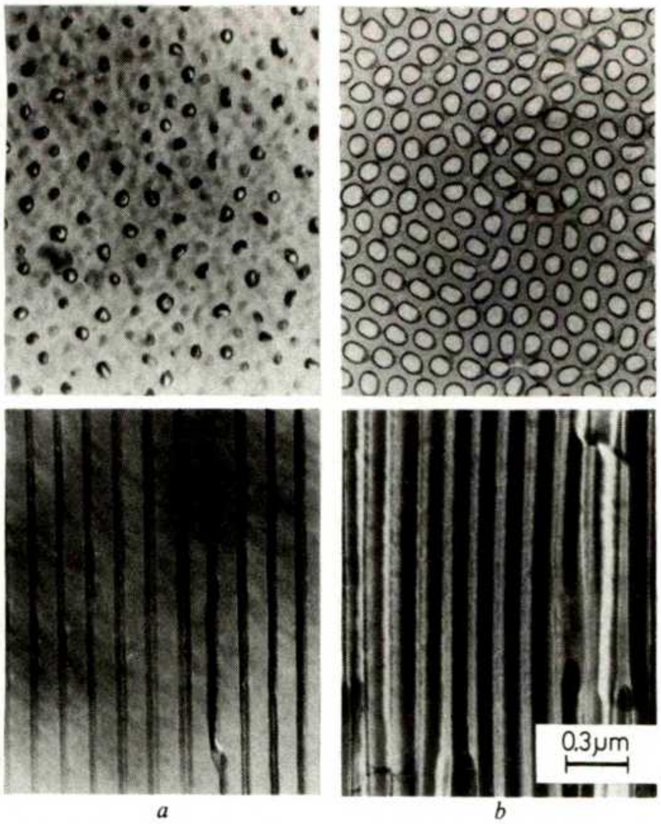


Fig. 1 Replica electron micrographs, showing the increase of pore diameters of the porous film with elapse of open-circuit dissolution time: *a*, as-anodized film formed on electropolished 99.99% aluminium at 3.0×10^{-2} A cm^{-2} in 0.5 M oxalic acid at 303 K for 60 min; *b*, the film dissolved for 140 min in 1M sulphuric acid at 313 K. Upper micrographs are the top surfaces and lower micrographs are the fractured sections of each film, respectively.

Figure 1 shows the open-circuit dissolution of the oxalic acid film in 1M sulphuric acid. The electron micrographs show that the pore widening occurs uniformly throughout the film[5]. Thus, it seems possible to determine the distribution of incorporated impurities across the pore walls by varying the dissolution period.

Figure 2 shows the variation of colour intensity of the film in terms of the integration of absorbance in the visible range (400–700 nm) by curve *A*, and PL intensity of the film with 310 nm excitation by curve *B* against the dissolution time T_d or the distance X from the surface of the pore wall. Curve *A* shows that the film decolours gradually in the first 90 min (0–225 Å), and then more rapidly during the next 90 min (225–450 Å). The yellow colour of the film disappears near this point and the film becomes semi-transparent. The distribution of the coloured materials calculated from curve *A* by differentiation with the amount of the dissolved oxide measured by gravimetry is shown in Fig. 3 *A*. The coloured material is located predominantly in the outer region of the cell walls (a, 0–225 Å); decreasing steeply in the next middle region (b, 225–450 Å) and none in the inner region (c, 450–540) Å). The distribution is shown in Fig. 3*B* together with the cell dimensions measured from the micrographs. This profile indicates that the hypothesis[6,7] of a uniform distribution of coloured materials in self-colouring films is no longer valid. In addition, this profile is in general agreement with that reported[4] for oxalate ions in the film calculated from the carbon content.

Curve *B* in Fig. 2 reveals that the PL centres are located predominantly in the middle region (b) of the pore walls. Note that the films dissolved for up to 90 min show relatively lower PL intensity, though the middle region still existed. This indicates that the outer coloured region (a) of the film tends to disturb the transmittance of excitation light, and emission light from the middle region (b).

Figure 4 shows the variation of the emission spectra with dissolution time when the film is excited by 310 nm light. The emission peak of the film shifts towards the shorter wavelength (470–435 nm) after a period of dissolution. Such a change became pronounced near the point ($T_d = 90$ min, $x = 225$ Å) when the outer coloured region (a) was dissolved. Furthermore, greenish-yellow phosphorescence with a peak at 560 nm with 360 nm excitation was observed in the coloured film containing region (a), but not in the inner film containing only region (b) and/or region (c). Therefore, luminescent centres in the outer region (a) and the inner regions (b, c) are clearly different.

In accounting for the appearance of these phenomena, there must be some conversion of aluminium–oxalate complex incorporated in Al_2O_3 matrix because the stoichiometric trioxalatoaluminate complex shows neither a yellow colour nor PL or ESR. A definite mechanism of the postulated conversion process has not yet been clarified; however, recent work[8] has produced a plausible model for the mechanism of successive transformation of aluminium–carboxylate complex into PL centres and coloured materials. According to this model, the carboxylate ions incorporated in the barrier oxide layer are subjected to decomposition followed by eventual polymerization. These processes are supposed to be induced by the collision of energetic (~2–5 eV, from EL spectra) electrons[9] that are injected into the conduction band of the oxide from the electrolyte at the oxide–electrolyte interface and have acquired such high kinetic energy due to the high electric field near 10^7 V cm^{-1} set up inside the oxide, or to be induced by Joule's heat. Thus, the carboxylate ions are expected to transform into intermediate compounds (luminescent centres) and finally, into coloured materials—mostly tarry products dispersed. This

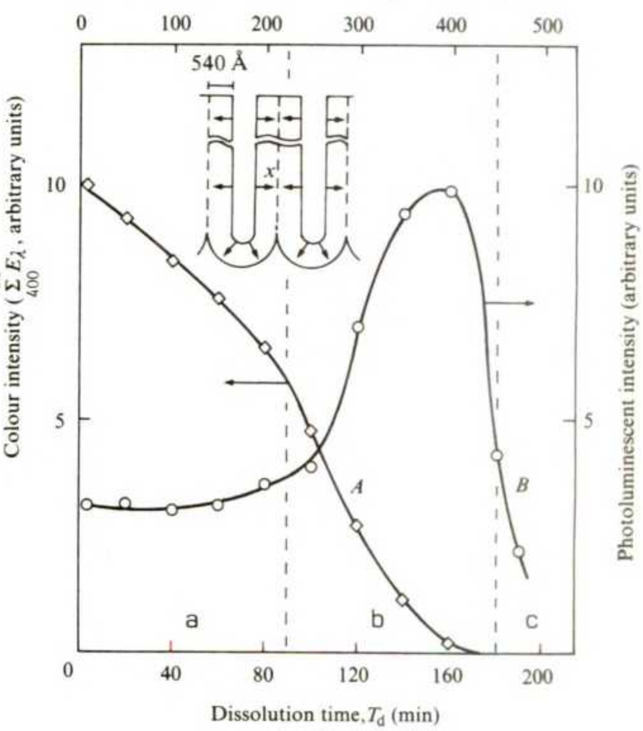


Fig. 2 Variations of the colour intensity of the film in terms of the integration of absorbance in the visible range (400–700 nm) (*A*) and PL intensity of the film with 310 nm excitation (*B*) against the time (T_d) of open-circuit dissolution or the distance (x) from the surface of the pore wall determined from mean dissolution rate of the film. The films were formed anodically on electropolished 99.99% aluminium at 3.0×10^{-2} A cm^{-2} in 0.5 M oxalic acid at 303 K for 60 min. After isolating the films from the metals, they were dissolved in 1 M sulphuric acid on open-circuit at 313 K for various times.

model is verified[10] by the thermosynthetic simulation of the anodic colouring process where the self-coloured alumina prepared by baking a mixture of hydrated $Al(OH)_3$ gel and carboxylic acid (molecular ratio 2 : 3) showed almost the same colour, PL and ESR characteristics as those of the anodically coloured alumina.

One possible explanation for the observed distributions of coloured materials and PL centres is given by the non-linear gradient of electric field across the barrier layer. The field strength, *F*, should increase between the cell base (wx) and the pore base (yz) (Fig. 3 *B*), as suggested by Hoar and Mott[11]. We can estimate the field strength in the oxide using the values given by Harkness and Young[12],

$$j = j_0 \exp -(w - qaF)/kT$$

where, $j_0 = 2.24 \times 10^7$ A cm^{-2}, $w = 1.3$ eV, $q = 3\,e$, $a = 2.95$ Å. In the present case, the field *F*, at the cell base corresponding to a current density of 3.0×10^{-2} A cm^{-2} at 303 K is then 8.96×10^6 V cm^{-1}. At the pore base, *F*, corresponding to a current density of about 3×10^{-1} A cm^{-2} at 303 K is then 9.60×10^6 V cm^{-1} because the pore base area is about one-tenth of cell base area. Thus, the field of outer region of the barrier layer is larger by ~7% than the inner region. In the outer region (a) of the barrier layer where the electric field is higher, the incorporations of oxalate ions and their conversions into PL centres or coloured materials should become more pronounced due to the high electric field. At steady state during anodization, the field-assisted dissolution[11,13] is balanced by the oxide growth at the cell base causing the continuous process of the anion incorporation and their conversion, and the transformation of the barrier layer to the cell wall of the porous layer. This would explain the development of distributions of the coloured materials and PL centres in the cell walls. However, the role of incorporated impurities on electric field cannot be neglected: an increase in the field has been shown by Dell'Oca and Young[14] whereas a decrease has been suggested by Thompson *et al.*[15]—this problem is still under discussion.

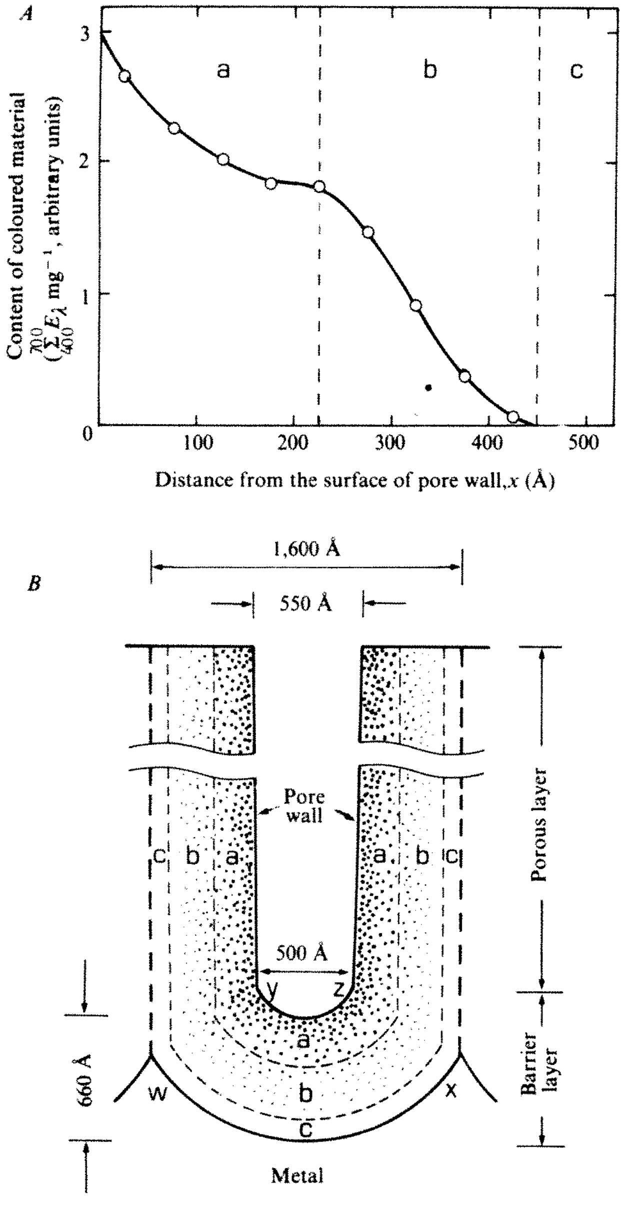


Fig. 3 *A*, The distribution of coloured materials in the porous alumina film formed at 3.0×10^{-2} A cm^{-2} in 0.5 M oxalic acid at 303 K, which is calculated from curve *A* in Fig. 2 by differentiation with the amount of dissolved oxide measured by gravimetry. The coloured materials are located predominantly in the outer region (a, 0–225 Å); decreasing steeply in the next middle region (b, 225–450 Å) and none in the inner region (c, 450–540 Å). *B*, The distribution of coloured materials in the film together with the cell dimensions measured from the electron micrographs.

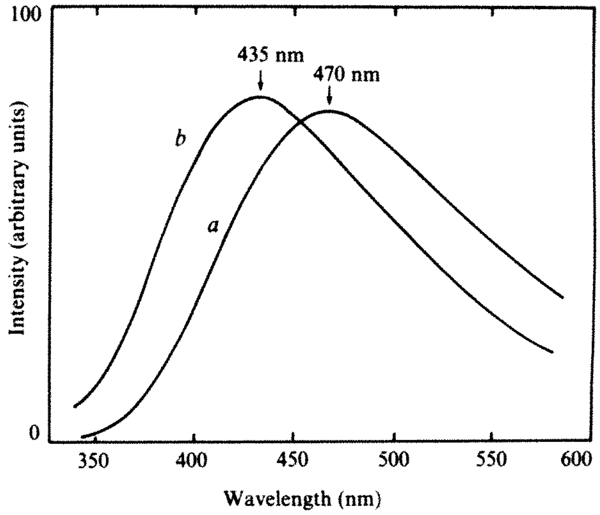


Fig. 4 The emission spectra of the films at different dissolution time when excited with 310 nm: *a*, as-anodized film ($T_d = 0$ min); *b*, the film dissolved for 160 min in 1 M sulphuric acid on open-circuit at 313 K.

Coloured materials and PL centres have different properties as a function of depth of the pore walls of the oxalic acid film. These results suggest that the incorporated oxalate ions which are present in various states within the oxide are converted during anodization. This may be attributed to the inhomogeneous distribution of the electric field across the oxide.

We thank Dr Y. Fukuda for discussions and The Mitsubishi Foundations, Tokyo for financial support.

Received 11 August; accepted 1 December 1980.

1. Tajima, S. *Electrochimica Acta* **22,** 995–1011 (1977).
2. Pullen, N. D. *J. Electrodep. tech. Soc.* **15,** 69–74 (1939).
3. Tajima, S. in *Advances in Corrosion Science and Techology* Vol. 1 (eds Fontana, M. G. & Staehle, R. W.) 229–362 (Plenum, New York, 1970).
4. Fukuda, Y. *J. chem. Soc. Jap.* 1868–1875 (1974).
5. Nagayama, M., Tamura, K. & Takahashi, H. *Corrosion Sci.* **10,** 617–627 (1970).
6. Sauter, W., Ibe, G. & Meier, J. *Aluminium* **50,** 143–149 (1974).
7. Almendinger, M. *Rev. Suisse d'Aluminium* 143–147 (1975).
8. Fukazawa, H., Baba, N., Yoshino, T. & Mizuki, I. *J. Metal Finish. Soc. Japan* **28,** 369–373 (1977).
9. Shimizu, K. *Electrochimica Acta* **23,** 295–298 (1978).
10. Yamamoto, Y., Baba, N., Yoshino, T. & Tajima, S. *Ext. Abstr. 29th ISE Meet.*, Budapest, 1014–1016 (1978).
11. Hoar, T. P. & Mott, N. F. *J. phys. Chem. Solids* **9,** 97–99 (1959).
12. Harkness, A. C. & Young, L. *Can. J. Chem.* **44,** 2409–2413 (1966).
13. O'Sullivan, J. P. & Wood, G. C. *Proc. R. Soc.* **A317,** 511–543 (1970).
14. Dell'Oca, D. J. & Young, L. *J. electrochem. Soc.* **117,** 1548–1511 (1970).
15. Thompson, G. E., Furneaux, R. C., Wood, G. C., Richardson, J. A. & Goode, J. S. *Nature* **272,** 433–435 (1978).

Trace element incorporation into growing augite phenocryst

N. Shimizu

Center for Geoalchemy, Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Since the introduction[1] to geochemistry of crystallization theories developed in metallurgy, it has been thought that the effects of kinetic parmeters such as the rate of crystal growth and the rate of diffusive transport of elements in the melt could significantly affect the partition of trace elements between mineral and magma. This conclusion has not, however, been demonstrated unequivocally, mainly because the extent of variations of equilibrium partition coefficients with *T*, *P* and chemistry is not well known. Henderson and Williams[2] found a correlation between morphology and apparent partition coefficient of uranium between olivine and basaltic melt. Data on diffusivity of elements (see ref. 3) and on crystal growth kinetics together with the development of secondary ion mass spectrometry (SIMS) techniques, enable crystal zoning to be studied with respect to trace elements and the kinetic effects to be evaluated quantitatively. Trace element zoning (or lack of it) of minerals is important, as many kinetic-based crystallization models predict trace element zoning in crystals. This letter presents trace element data on a sector-zoned augite phenocryst based on the ion-probe spot analyses. The fact that the slower-growing prism sector [100] is enriched in both compatible and incompatible elements relative to the faster-growing basal sector [$\bar{1}$11] strongly supports Dowty's[4] model for crystal growth involving preferential adsorption of elements onto growing crystal faces in proportion to the charge/size ratio of the elements.

To enhance possible kinetic effects, sector-zoned augite phenocrysts were used, as they present a clear case of disequilibrium crystallization. In a sample of alkali olivine basalt lava (collected by M. Hill from the western foot of Bald Mountain, near Truckee, California) Ti-augite phenocrysts up to 1 × 2 mm in size with very well developed sector zoning occur with olivine and plagioclase phenocrysts in a holocrystalline groundmass. The particular phenocryst studied here has dimensions 0.5 × 1 mm, and a section of this crystal is sketched in Fig. 1. The dark-coloured [100] sector and the pale [$\bar{1}$11] sector share a sharp boundary.

The analytical data were obtained with the MIT–Brown–Harvard Cameca IMS 3F ion probe; the techniques will be described in detail elsewhere, but are similar in principle to those described previously[5]; in particular, energy filtering was used to eliminate molecular ion interferences effectively. A beam of O^- ions with net energy of 13.1 keV was focused to a spot 5–8 μm in diameter on the gold-coated surface of a polished thin section. The secondary ion intensities were measured by an electron multiplier coupled with pulse-counting circuitry, and intensity ratios relative to Si were measured for Na, Mg, Al, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Sr and Zr. As shown previously[5], the relative intensities of most elements are proportional to their concentrations. Data are presented for some of these as relative intensities; absolute concentrations are reported for those elements for which precise standards and working curves are available.

Traverses consisting of a series of closely spaced points were made at various locations in the crystal (Fig. 1) to study core-to-rim zoning in individual sectors, and to study inter-sector relationships. Data from one of the traverses are shown in Fig. 2. Note that significant differences exist between the two sectors with respect to most of the elements. Also, the [100] sector is homogeneous from core to rim for all elements except Cr. Another traverse shows that the [$\bar{1}$11] sector is also homogeneous from core to rim and that the same outward decrease in Cr is observed. On the basis of Cr variation, this crystal consists of a central region or plateau of high Cr, a marginal region or plain of low Cr and a transition zone in between. These results show that the crystal grew in different directions simultaneously.

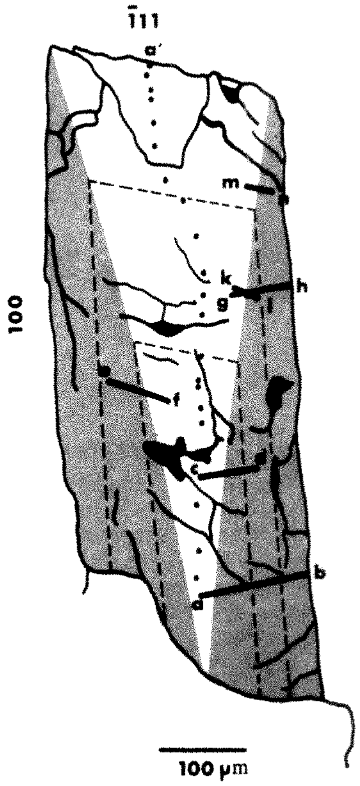


Fig. 1 A crystal of sector-zoned augite. Traverses are labelled a–b, c–d, and so on. A series of points connecting a–a' forms a core to rim traverse of the [$\bar{1}$11] sector. The analytical data along the traverse a–b are shown in Fig. 2. Dashed lines mark approximate locations of the central high Cr plateau, marginal low Cr plain and the transition zone in between (see text).

The most surprising observation is that the slower-growing prism sector [100] is enriched in both compatible (for example, Cr and Sc) and incompatible (for example Ti and Zr) elements relative to the faster-growing basal sector [$\bar{1}$11], even though both sectors are internally homogeneous.

If surface equilibrium was approached during trace element partitioning at the crystal–melt interface, the effect of the faster growth of the [$\bar{1}$11] sector would be to deplete that face in Cr and Sc and enrich it in Al, Ti, Zr, and so on relative to the [100] face. Also, if the overall crystal growth rate was sufficiently rapid relative to the rate of diffusive transport of elements in the magma, an outward decrease of Sc and Cr should be accompanied by an increase of Al, Ti and Zr. Details of these model cases are discussed elsewhere[1]. Neither of these suppositions, which are logical consequences of a simple interplay between the crystal growth rate and diffusion rate, is consistent with the observed data.

An implicit assumption in the crystallization models[6–8] is that interface kinetics are sufficiently rapid compared with diffusivity and the crystal growth, thereby ensuring equilibrium at the crystal–melt interface. However, crystal growth in many glass-forming systems is limited by interface kinetics rather than the rate of removal of latent heat (see, for example, refs 9, 10) and there is evidence that crystallization of silicate minerals from magma is rate-limited by a similar process (see, for example, refs 11–13). Interface kinetics is a loose term which includes processes such as adsorption and desorption, dissociation and chemical reaction, surface and edge migration[14], and we need to consider

0028–0836/81/070575–03$01.00

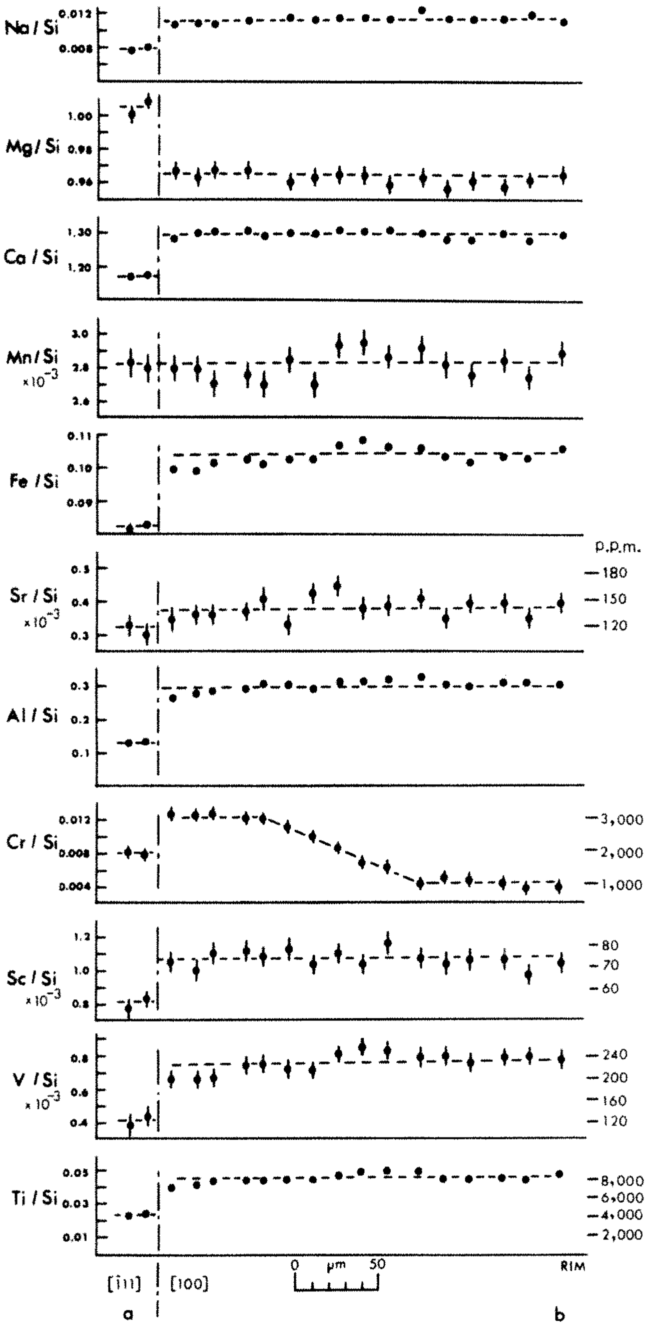


Fig. 2 Analytical data along the traverse a–b. The relative intensities of elements against Si are shown. The concentrations calculated are shown at the right-hand side of the diagram. The error bars represent 2σ counting error. The sector boundary is shown as a vertical broken line.

specific physical processes operating during crystal growth with special attention to the mechanisms of trace element incorporation.

In this context, the model of formation of sector-zoned augite[4,15] has an important bearing on trace element incorporation. The present discussion is essentially an extension of this model to trace elements. The basis of the model is that on the surface of growing augite there is a certain number of cation protosites (which are future M sites), defined by a concentration of negative charges derived from partially bound oxygen atoms, to which cations can be attracted (or adsorbed). In the direction perpendicular to (100) of augite, for instance, there are alternate 'layers' of SiO_3-chains and M_1–M_2 cations; when a new SiO_3 chain is attached, there are several protosites available on the surface to which cations may be adsorbed. If the attachment of SiO_3-chains is rapid, the adsorbed cations are left entrapped; if the attachment is slow, however, the adsorbed elements have time to be desorbed and finally to reach an equilibrium distribution for the given conditions. Thus, the incorporation of trace elements into a growing augite crystal is controlled by two kinetic processes: (1) the attachment kinetics of SiO_3-chains; and (2) the adsorption/desorption kinetics of cations in the protosites. Therefore, for the case of moderately rapid SiO_3 attachment, equilibrium is not expected to be achieved and the degree of adsorption/desorption is expected to be a function of how strongly the cations are bonded to the protosites; the stronger the adsorption, the higher the activation energy of desorption and the slower the desorption, the higher the retention of the adsorbed cations.

Obviously, however, this interrelation of two kinetic processes alone does not account for the fact that the slower-growing [100] sector is enriched in trace elements relative to the faster-growing $[\bar{1}11]$ sector (Fig. 2). A crucial factor here is the number of protosites available in each growth direction. For simplicity, only elements entering the M_1 site, and only the 3/6 and 4/6 types of protosites (see ref. 4) will be considered here. If the growth rate is sufficiently rapid in all directions, the number of atoms adsorbed and then left entrapped will be proportional to the number of available protosites in each direction. Therefore, the [100] sector is expected to take up four times as many atoms as the $[\bar{1}11]$ sector, and the enrichment should be the same for all elements. The actual enrichments in [100] over $[\bar{1}11]$ observed in this study, however, vary between 1 and 2 for the different elements. This suggests that different elements were desorbed from a given sector to different degrees and that desorption was more extensive from the slower-growing [100] sector than the faster $[\bar{1}11]$ sector.

The different degrees of desorption among elements must reflect a difference in the activation energy of desorption (equivalent to a difference in adsorption energy). Figure 3 shows that the preferential enrichment of elements in [100] relative to $[\bar{1}11]$ is proportional to the ratio of the ionic charge of the element divided by the ionic radius squared. The ordinate is a measure of the degree of desorption while the abscissa represents the strength of the Coulomb interaction between a given cation and the protosite. Assuming that the initial enrichment or 'adsorption' factor equals a maximum value of 4 for all elements (based on the relative abundance of protosites), Fig. 3 shows that the stronger the Coulomb interaction, the higher the retention of cations (or the lower the degree of desorption). Dowty[4] suggested that the cations which will be bound most firmly to the protosites are those which have high

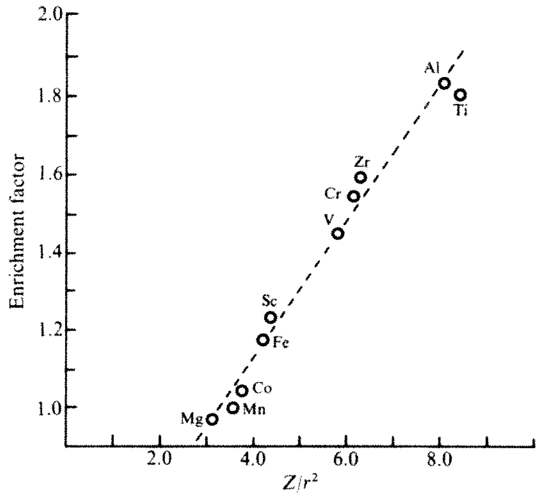


Fig. 3 Preferential enrichment of elements in the [100] sector relative to $[\bar{1}11]$ sector as a function of the ratio of the ionic charge of the element divided by the ionic radius squared. The ionic radii are taken from ref. 16. The correlation coefficient of the points is 0.977.

charges and small size. Figure 3 demonstrates that this relationship holds for 10 elements entering the M_1 site; the Coulomb interaction between cation and protosite is thus the controlling mechanism in this process.

As crystallization conditions (such as the temperature of the magma, the degree of supercooling and the cooling rate of the system) vary from one magma body to another, the relative kinetics of SiO_3 attachment and cation adsorption/desorption are expected to vary, thereby producing different inter-sector element ratios. Conversely, such cation distributions can be used to study different kinetic effects among sector-zoned augites from different areas representing a range of crystallization conditions. A more detailed account of the data presented here, including absolute concentrations of all elements, apparent partition coefficients, and comparison among sectors other than [100] and [$\bar{1}$11], will be published elsewhere.

The present data strongly support the theoretical model of Nakamura[15] and Dowty[4]. The data are not consistent with models such as those given in refs 6–8 which involve an interplay of crystal growth rate and element diffusion rates in the melt adjacent to growing crystals.

I thank S. R. Hart for discussions and comments, also M. Hill for the sample used in this study, K. D. Burrhus for technical assistance, and D. S. Hall for artwork. The work is supported by the NSF grants EAR 7922049 and EAR 8006642.

Received 26 September; accepted 26 November 1980.

1. Albarède, F. & Bottinga, Y. *Geochim. cosmochim. Acta* **36,** 141 (1972).
2. Henderson, P. & Williams, C. T. *Phys. Chem. Earth* **11,** 191 (1979).
3. Hofmann, A. W. in *Physics of Magmatic Processes* (ed. Hargraves, R. B.) 385 (Princeton University Press 1980).
4. Dowty, E. *Am. Miner.* **61,** 460 (1976).
5. Shimizu, N., Semet, M. P. & Allegre, C. J. *Geochim. cosmochim. Acta* **42,** 1321 (1978).
6. Tiller, W. A., Jackson, K. A., Rutter, J. W. & Chalmers, B. *Acta met.* **1,** 428 (1953).
7. Smith, V. G., Tiller, W. A. & Rutter, J. W. *Can. J. Phys.* **33,** 723 (1955).
8. Burton, J. A., Prim, R. C. & Slichter, W. P. *J. Chem. Phys.* **21,** 1987 (1973).
9. Meiling, G. S. & Uhlmann, D. R. in *Crystal Growth* (ed. Peiser, H. S.) 747 (Pergamon, Oxford, 1967).
10. Vergamo, P. J. & Uhlmann, D. R. *Phys. Chem. Glasses* **11,** 30 (1970).
11. Kirkpatrick, R. J. *Am. Miner.* **60,** 798 (1975).
12. Kirkpatrick, R. J., Robinson, G. R. & Hays, J. F. *J. geophys. Res.* **81,** 5715 (1976).
13. Klein, L. & Uhlmann, D. R. *J. geophys. Res.* **79,** 4869 (1974).
14. Ghez, R. & Lew, J. S. *J. Crystal Growth* **20,** 273 (1973).
15. Nakamura, Y. *Am. Miner.* **58,** 986 (1973).
16. Whittaker, E. J. W. & Muntus, R. *Geochim. cosmochim. Acta* **34,** 945 (1970).

Plagiogranites from the ocean crust and ophiolites

D. T. Aldiss*

Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK

Rock types similar to those found in ophiolites have been recovered from the floors of ocean basins[1–3]. This is critical evidence for the theory that ophiolites represent the ocean crust[4,5]. Aumento[6] has described diorites, quartz diorites and trondhjemites dredged from the Mid-Atlantic Ridge (MAR) near 45°N, and noted that these were the first rocks analogous to the low-K acid differentiates in ophiolites (collectively referred to as 'oceanic plagiogranites'[7,8]) to be recognized amongst those recovered from the modern ocean floor. Only a few plagiogranites from the ocean floor (ocean-crust plagiogranites) have since been described; these were dredged from widely spaced localities on the MAR[9–11], the Amani Plateau and the Kyushu–Palau Ridge (north-west Pacific)[12]. However, the Kyushu–Palau Ridge could be the product of destructive plate-margin magmatism[13], as are the potassic granites from the Aves Ridge (Caribbean)[14]. Plagiogranite xenoliths also occur in extrusives on Surtsey and Askja (Iceland)[15]. That only some of the petrological compositions of ophiolite plagiogranites are represented in the ocean crust is not surprising in view of the small sample of ocean-crust plagiogranites. It is less obvious, however, if ophiolites represent the ocean crust, why certain acidic rock types from the oceans have not been found in ophiolites. I show here that the difference between the two groups of plagiogranites is evidence that ophiolites do not represent all kinds of ocean crust and that a greater range of magma compositions and/or conditions of magmatic differentiation is represented by the ocean crust.

Considering that there have been ~500 DSDP boreholes and of the order of 10,000 oceanic dredge hauls, the number of ocean-crust plagiogranites seems very small even compared with the low abundance of plagiogranites in ophiolites[7,8,16,17]. Although the oceanic sample is presumably biased to represent shallow structural levels of the ocean crust, this small number of examples seems to indicate that plagiogranite forms an even smaller proportion of the ocean crust than of most ophiolites.

Plagiogranites from the ocean crust and those from ophiolites are generally similar in essential primary mineralogy (predominantly sodic plagioclase with quartz and hornblende), in secondary minerals and textures and in major element geochemistry (plagiogranites are characterized by $K_2O < 0.7\%$ and the more silicic are relatively depleted in Fe, Ti, P and K and more enriched in Na)[1,6–11,17].

However, there are differences in accessory mineralogy and alkali and trace element geochemistry between some ocean-crust plagiogranites and typical ophiolite plagiogranites. Except in some large (few kilometres diameter) plagiogranite stocks and aplites which intruded some ophiolites at a late stage in their development[17,18] and the Byne Hill (Ayrshire) trondhjemite[19], K-rich minerals (biotite and so on) are virtually absent from ophiolite plagiogranites[7,8,16,17]. I have found rare secondary

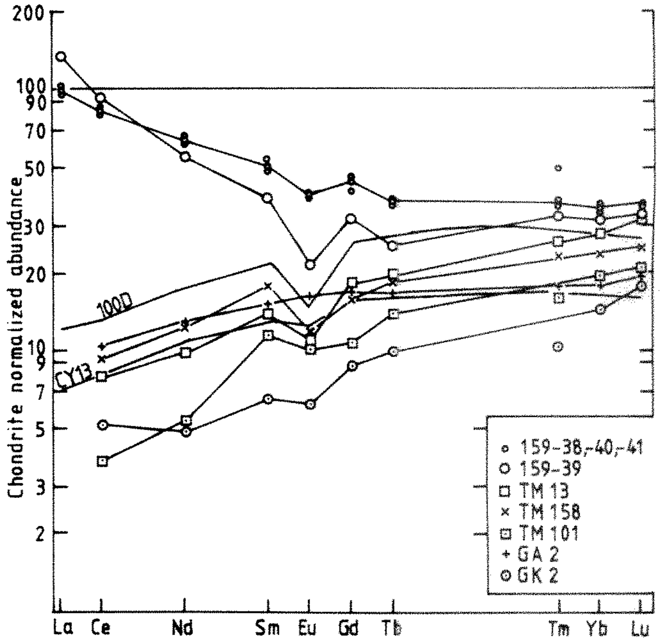


Fig. 1 Chondrite normalized rare earth profiles of some ophiolite and ocean crust plagiogranites. 159-38, 39, 40, 41: quartz diorites from the MAR at 45° N. TM13, 101, 158, GA2, GK2: tonalites and trondhjemites from the Troodos Massif, Cyprus[17]. CY13, 100D: Troodos plagiogranites[23]. Chondritic abundances after ref. 29.

---

* Present address: Overseas Division, Institute of Geological Sciences, Keyworth, Nottingham NG12 5GG, UK.

0028-0836/81/070577-02$01.00

Table 1 REE, Th and Ta content* of some ocean-crust plagiogranites

	AG159-38	AG159-39	AG159-40	AG159-41	ANTP125-4(c)	TM	OM
La	35.1	43.8	32.7	33.4	—	—	3.5
Ce	74.3	81.7	71.6	73.0	53.2	6.7	10.5
Nd	42.2	35.1	39.6	39.3	30.3	6.1	—
Sm	11.0	7.7	10.1	10.0	7.00	2.8	3.3
Eu	3.11	1.62	3.05	2.80	0.616	0.88	1.0
Gd	12.6	8.8	12.5	11.3	8.22	3.7	—
Tb	2.0	1.3	1.9	1.9	—	0.83	0.85
Tm	1.4	1.1	1.3	1.2	—	0.7	—
Yb	7.7	6.9	7.3	7.4	8.22	4.6	4.2
Lu	1.2	1.1	1.2	1.2	—	0.8	0.7
Th	4.5	8.7	3.8	2.3	—	0.6	—
Ta	3.2	3.2	2.8	3.5	—	1.0	—

AG159, Quartz diorites from the MAR at 45° N. ANTP 125-4(c), Quartz monzonite from Argo Fracture Zone[21]. TM, Average of five tonalites and trondhjemites from the Troodos Massif, Cyprus. OM, Average of two plagiogranites from the Semail ophiolite, Oman[8]. New data (for AG159 and TM) determined by instrumental neutron activation analysis at the Open University. Full analytical details are given in ref. 17.

* All values in p.p.m.

allanite in an epidote–quartz metamorphite from the Troodos Massif (Cyprus) but know of no other occurrence of this mineral in an ophiolite. However, a diorite from the MAR at 45°N has rare K-feldspar cores, minor biotite and allanite[6]. Elsewhere on the MAR, biotite dominates hornblende in a quartz diorite from 0°N (ref. 9) although plagiogranites from 22° S resemble typical ophiolite plagiogranites in every descriptive detail[10]. Plagiogranite xenoliths from Surtsey have minor K-feldspar and relatively high K_2O (ref. 15).

Some granitoids from the ocean-crust are not plagiogranites. A specimen of leucoquartz monzonite (ANTP 125-4) dredged from the Argo Fracture Zone, Central Indian Ridge[20,21], contains essential K-feldspar and also has high Rb, Nb and Y and low K/Rb, relative to ophiolite plagiogranites. Two aplites from the Argo Fracture Zone also have high Nb and Y (Table 2). The only similar rocks from any ophiolite are late alkalic aplites which intruded gabbros of Masirah Island, North Indian Ocean, and which are believed to have originated in continental crust[22].

Four of the quartz diorites from the MAR at 45°N were analysed for rare earth elements (REE), Th and Ta. They are light REE (LREE) enriched and also have a greater abundance of incompatible trace elements than representative ophiolite plagiogranites (Table 1, Fig. 1). The REE pattern[21] of ANTP 125-4 is similar to that of the quartz diorites although with a more pronounced Eu anomaly. Ophiolite plagiogranites are normally LREE depleted although they also have negative Eu anomalies[7,8,17,23] (Fig. 1). These contrasts in trace element content could be attributed to regional differences in either magma composition or magmatic differentiation. (Hydrothermal metamorphism, although an important petrogenetic factor in ophiolites and the ocean crust[24,25], is thought not to change the distribution of the REE[26,27]). Although basalts enriched in LREE and other large ion lithophile elements have been found on the MAR at 45°N, as well as the usual LREE depleted varieties[28], basic rocks associated with the Argo Fracture Zone granitoids are all LREE depleted[21]. Therefore primary magma composition need not be the controlling factor. Relative LREE-enrichment of ANTP 125-4 has been attributed to apatite fractionation[21] and more advanced differentiation could also account for its enrichment in other incompatible trace elements.

Table 2 Comparative trace element content of Argo Fracture Zone granitoids and ophiolite plagiogranites

	Quartz monzonite* (ANTP 125-4(c))	Aplite*	Aplite*	Average ophiolite plagiogranite†		
				$\bar{x}$	Range	*n*
K_2O (%)	3.27	0.07	—	0.21	0.02–0.68	56
Rb (p.p.m.)	89.5	—	—	<3	<1–5	56
Nb (p.p.m.)	29	25	≈20	<4	<1–14	49
Y (p.p.m.)	150	180	120	40	3–88	50
K/Rb	303	—	—		>500	12

*Argo Fracture Zone granitoid[20,21].

†Quartz diorites, tonalites, trondhjemites and aplites from the Troodos Massif, Cyprus, the Semail ophiolite, Oman and the Smartville Block, California[8,17].

It seems therefore that some ocean-crust plagiogranites cannot be distinguished from typical ophiolite plagiogranites and that their similarities are greater than their differences. This supports the contention that ophiolites represent ancient ocean crust. However, the mineralogical and chemical variety of granitoids discovered in the present-day ocean crust exceeds that found in ophiolites although the actual number of examples is much smaller, suggesting that a greater range of physical and/or chemical conditions pertained during the formation of the ocean crust. It also implies that ophiolites are a sample of a limited number of types of ocean crust. I look forward to descriptions of comparative material from a modern marginal basin or juvenile island arc.

I thank Dr F. Aumento for samples of quartz diorite, the NERC for support, Professor I. G. Gass for supervising my research studentship, Dr Phil Potts for help with REE analysis, and Drs R. G. Coleman and N.M.S. Rock for helpful criticisms.

Received 22 April; accepted 11 December 1980.

1. Aumento, F., Loncarevic, B. D. & Ross, D. I. *Phil. Trans. R. Soc.* A**268,** 623–650 (1971).
2. Cann, J. R. *Phil. Trans. R. Soc.* A**268,** 605–617 (1971).
3. Thompson, G. *Chem. Geol.* **12,** 99–111 (1973).
4. Gass, I. G. *Nature* **220,** 39–42 (1968).
5. Dewey, J. F. & Bird, J. M. *J. geophys. Res.* **76,** 3179–3206 (1971).
6. Aumento, F. *Science* **165,** 1112–1113 (1969).
7. Coleman, R. G. & Peterman, Z. E. *J. geophys. Res.* **80,** 1099–1108 (1975).
8. Coleman, R. G. & Donato, M. M. in *Trondhjemites, Dacites and Related Rocks* (ed. Barker, F.) (Elsevier, New York, 1979).
9. Bonatti, E., Honnorez, J. & Ferrara, G. *Earth planet. Sci. Lett.* **9,** 247–256 (1970).
10. Christensen, N. I. *Earth planet. Sci. Lett.* **36,** 297–300 (1977).
11. Miyashiro, A. *Earth planet. Sci. Lett.* **7,** 361–365 (1970).
12. Ishizaka, K. & Yanagi, T. *Earth planet. Sci. Lett.* **27,** 371–377 (1975).
13. Uyeda, S. *Bull. geol. Soc. Am.* **85,** 1159–1170 (1974).
14. Walker, B. M., Vogel, T. A. & Ehrlich, R. *Earth planet. Sci. Lett.* **15,** 133–139 (1972).
15. Sigurdsson, H. *Geol. mag.* **105,** 440–453 (1968); *Nature* **269,** 25–28 (1977).
16. Coleman, R. G. *Ophiolites* (Springer, Berlin, 1977).
17. Aldiss, D. T. thesis, Open Univ. (1978).
18. Alleman, F. & Peters, T. *Eclog. geol. Helv.* **65,** 657–697 (1972).
19. Bloxham, T. W. *Trans. R. Soc. Edinb.* **68,** 105–122 (1968).
20. Engel, C. G. & Fisher, R. L. *Bull. geol. Soc. Am.* **86,** 1553–1578 (1975).
21. Hedge, C. E., Futa, K., Engel, C. G. & Fisher, R. L. *Contr. Miner. Petrol.* **68,** 373–376 (1979).
22. Abbotts, I. L. *Geol. Mag.* **115,** 415–425 (1978).
23. Kay, R. W. & Senechal, R. G. *J. geophys. Res.* **81,** 964–970 (1976).
24. Spooner, E. T. C. & Fyfe, W. S. *Contr. Miner. Petrol.* **42,** 287–304 (1973).
25. Gass, I. G. *Geol. Soc. Lond. Spec. Publ.* **7,** 72–76 (1977).
26. Pearce, J. A. *Tectonophysics* **25,** 41–67 (1975).
27. Wood, D. A., Gibson, I. L. & Thompson, R. N. *Contr. Miner. Petrol.* **55,** 241–254 (1976).
28. Frey, F. A., Bryan, W. B. & Thompson, G. *J. geophys. Res.* **79,** 5507–5527 (1974).
29. Nakamura, N. *Geochim. cosmochim. Acta* **38,** 757–775 (1974).

Lower Cretaceous tides revealed by cross-bedding with mud drapes

J. R. L. Allen

Department of Geology, The University, Reading RG6 2AB, UK

The sand-dominated Folkestone Beds, widely exposed around the Weald of south-east England, are of Upper Aptian to Lower Albian (Lower Cretaceous) age and range in thickness up to 84 m (refs 1–3). Their content of ammonites, bivalves and sponges indicates a marine origin, while their relative coarseness and the incorporation of drifted timber point to a shallow water deposition near land. One striking sedimentary facies is formed by medium to coarse-grained cross-bedded sands in which several foreset and/or bottomset surfaces support mud drapes. Although it was suggested[4] that this facies may have been formed by the storm-induced movement of sand waves, I now consider that the facies is best explained by the migration of sand waves under the influence of strongly asymmetrical tides, and that the mud drapes are formed at slack water. A model of this process has been described elsewhere[5] and I now describe measurements of a section which supports my earlier conjecture[6] that the separation of the mud drapes is systematically related to the phase of the spring–neap tidal cycle at the time of deposition. Although criteria for recognizing tidal deposits have long been available it has not previously been possible to analyse a series of ancient sediments with a view to characterizing the tidal regime.

Measurements were made of a 22.9-m long portion of a typical cross-bedding set exposed on a smooth vertical face of the Pendean Sand Pit near Midhurst, Sussex (SU 890 200; operated by Hall Aggregates (South Coast) Ltd). The set is bounded above and below by nearly horizontal surfaces corresponding with erosion levels. The set shows groups of sandy foresets–bottomsets (taken to correspond with episodes of sand transport and deposition) alternating laterally with much thinner mud drapes (corresponding with episodes of mud deposition). The mud drapes are rarely thicker than a few millimetres, although locally several may be piled on each other to make a thick mud layer. Many of the mud drapes have one or—rarely—more internal silty or sandy partings. The extent of a drape as revealed on the quarry face is variable, Fig. 1 showing a schematic profile in the cross-bedding set.

The measured set is ~2 m thick. The foreset azimuth is nowhere greatly different from the trend of the face (15°–195°) so that the measured thickness of each sand deposit (group of foresets) is very nearly its true thickness in the direction of the tidal stream at deposition. The 22.9-m section includes 65 groups of sandy foresets and 66 intervening mud drapes of different extent (classified as short, medium and long, Fig. 1). The apparent thickness (d) of the successive groups of sandy foresets is shown in Fig. 2, together with the estimated extent of the intervening mud drapes.

The marked periodic variation of the thickness d in the records of Fig. 2 accounts for the way in which field observation shows the mud drapes to be bunched or bundled together—bunching occurs where the sandy foresets are thinnest. The section I have measured includes four complete cycles (numbered 2–5) (including 32, 28, 16 and 30 episodes) and parts of two other cycles (cycles 1 and 6 in Fig. 2) including at least 18 and 6 episodes respectively. The complete cycles are consistently asymmetrical, thickness increasing more quickly than it decreases. On descending limbs there usually occur alternately thick and thin sand deposits (for example, episodes 64–78, 84–94, 100–124 all inclusive). A further 28 cycles (24 counted, 4 estimated) lie in the contiguous downstream extension of the same set. The longer and thicker drapes (A–E on Fig. 2) tend to be associated with, or lie just upstream from, the thinner sand deposits.

It has been proposed[5] that the formation of this facies was dependent on a subtidal sandwave influenced by a strongly asymmetrical tide. The variation of stream speed in such a cycle (of length T) is shown in Fig. 3 Asymmetrical tidal patterns may arise in shallow water because of the distortion of the tidal wave, or because of the presence of a steady current, which may be caused in several ways, or both.

The instantaneous rate of mud deposition (mass per unit area of bed surface) from such a stream may be written[7] as

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \sigma C W (1 - U^2/U_{\mathrm{cdm}}^2); \; U \leq U_{\mathrm{cdm}} \qquad (1)$$

where m is the mass per unit area, t is time, σ is the density of mud solids, C the fractional volume concentration of mud in the water, W the falling velocity of mud particles, U the vertical average of current speed, and U_{cdm} the speed above which mud deposition does not occur evaluated at the same roughness as U.

Figure 3 also shows the two critical stream speeds on which our model depends. These are U_{cdm}, the critical speed below which mud deposition can occur, and U_{ces}, the critical speed above which bedload sand movement occurs. U_{ces} is greater than U_{cdm}.

Over a tidal cycle (Fig. 3), mud will accumulate during the intervals $(t_5 - t_4)$ and $(T - t_6 + t_1)$, where T is the tidal period. But as the tidal streams are asymmetrical, but at times exceed the threshold for bedload sand movement, U_{ces}, between t_5 and t_6, the two accumulations should tend to appear as one in the field, with the strong possibility that a silty or sandy parting may be formed, due either to the mobilization of grains at the sand-wave crest or continued fall-out of bed material during the peak stream following t_5. This explains the partings seen in the field within many drapes. Moreover, as U_{cdm} is of the order of 0.05–0.10 m s^{-1}, the mere presence of the drapes implies that the tide was not significantly rotary and that the cross-bedding set therefore accumulated in restricted waters, either near shore, in an estuary, or within a strait. Integration of equation (1) over the relevant parts of the tidal cycle shows that the drapes should be thickest and longest when the peak tidal currents are relatively gentle (other factors remaining constant), for $(t_5 - t_4)$ and $(T - t_6 + t_1)$ are then maximized. This tends to occur when the sand deposits take their smallest apparent streamwise thickness.

The spatially-averaged instantaneous bedload sediment transport rate beneath a tidal stream fashioning sand waves is approximately described kinematically by

$$J = k(U - U_{\mathrm{ces}})^3, \qquad U \geq U_{\mathrm{ces}} \qquad (2)$$

in which J is the mass carried down a lane of unit width in unit time, and k a dimensional constant set by sediment and flow characteristics[5,8,9]. Hence the true streamwise thickness of the sand deposit formed during one tidal cycle (Fig. 3) is

$$d = \frac{2k}{H\gamma} \int_{t_2}^{t_3} [U(t) - U_{\mathrm{ces}}]^3 \, \mathrm{d}t, \qquad U \geq U_{\mathrm{ces}} \qquad (3)$$

in which H is the sand-wave height, and γ the sand bulk density.

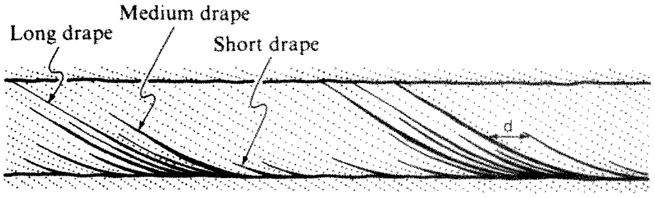


Fig. 1 Schematic representation of cross-bedding set with bundles of mud drapes (individual drapes distinguished as short, medium or long).

0028–0836/81/070579–03$01.00

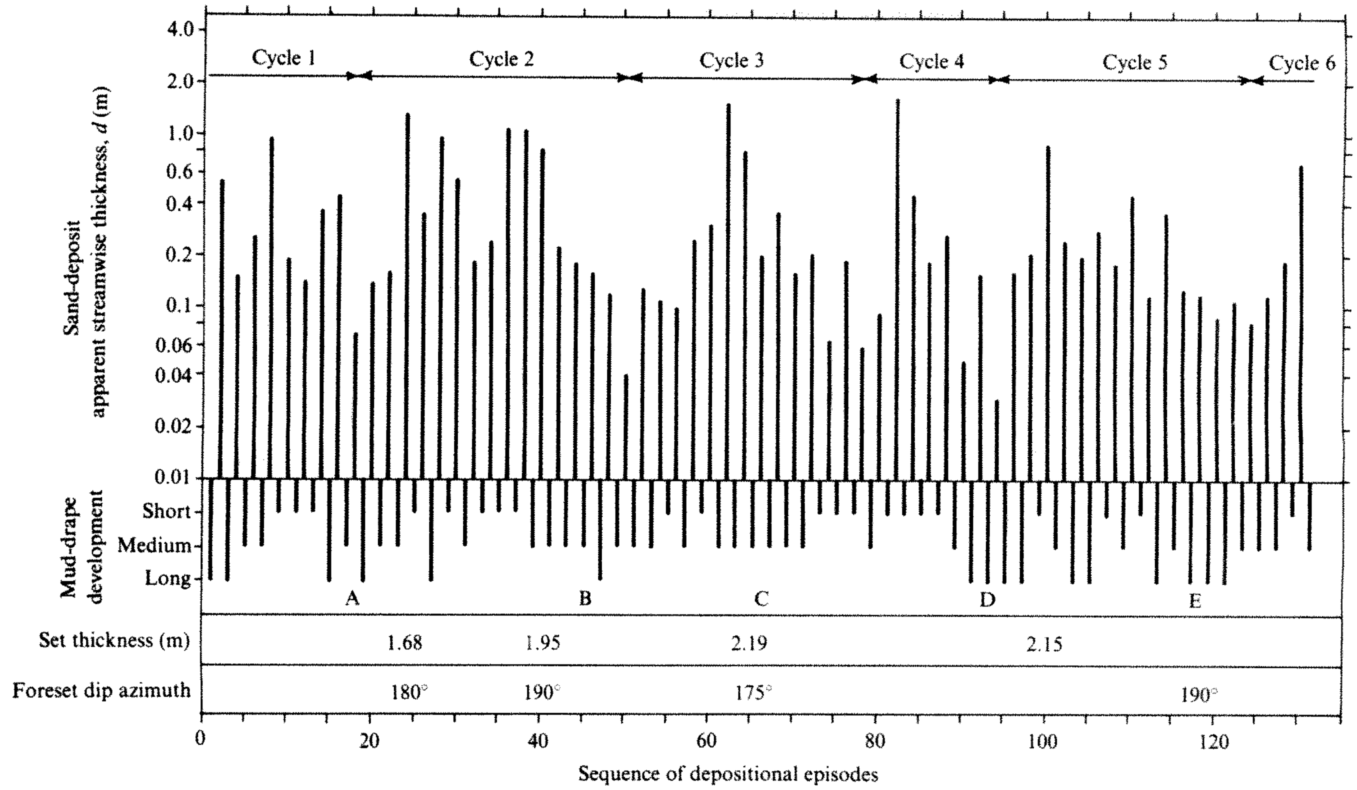


Fig. 2 Sand-deposit apparent streamwise thickness (d) and mud-drape development as a function of time (depositional episodes) in a cross-bedding set at Pendean Sand Pit near Midhurst.

Because J is a steeply increasing function of the speed difference, a small change in the peak speed of the tidal current should create a large change in the sand-deposit streamwise thickness. Applied to the Pendean thickness record (Fig. 2), equation (3) points to a substantially regular variation in the maximum speed of the tidal currents that deposited the set, the weaker peak streams, yielding the thinner layers, occurring at about the same time as, or a little after, the interval when most mud settled out. The observed antisymmetrical behaviour of the mud and sand deposits is thus fully explained by the model, although many factors, such as a change in mud concentration, the growth of wind waves and a resultant increase in bottom stress and inhibition of mud deposition, or the superimposition of storm currents on the tidal stream, may have disturbed its theoretical regularity. Could the slight asymmetry of the sand thickness cycle be related to bed-hardening consequent on extensive mud deposition and ageing[10] as the peak tidal currents declined towards the neaps?

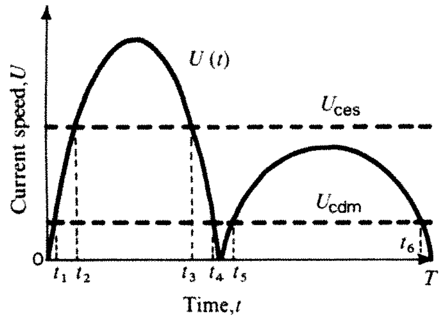


Fig. 3 Schematic time–speed pattern for tidal currents during one tidal cycle of the general kind thought to generate strongly asymmetrical sand waves (note that bedload transport is restricted to one half of the cycle only in the specific instance graphed; in some cycles U_{ces} may be exceeded slightly by the weaker stream).

The above factors may also account for the slight lack of uniformity in the period of sand thickness variation, for the suppression of just a few mud drapes would drastically cut the number of detectable depositional episodes (? cycle 4). Alternatively, during some neap tides the peak stream may never have equalled U_{ces} in speed, when no sand deposit would have resulted. The observed preferred period of 28–32 episodes strongly suggests that the thickness cycle records a spring–neap tidal periodicity (thick layers ≡ spring tides; thin layers ≡ neaps) involving tides of the diurnal type[11], a bidiurnal inequality (unknown in present-day tides) being strongly suggested by the regularly declining sequences of alternately thick and thin foreset groups (see, for example, episodes 64–78). Alternatively, the tide could have been semidiurnal, but with many flows failing to reach the sand movement threshold. This interpretation is thought less likely, because in association with the thinner sand deposits there are inadequate numbers of thick mud drapes including more than one sand–silt parting. That the period inferred from the Lower Cretaceous Folkestone Beds is a little larger than that measured today is consistent with palaeontological and other evidence[12–14] for a gradual and slight reduction through time in the number of days in the year. Sedimentological evidence for this change has not previously been available.

Careful analysis of the pattern of lithological change (or reactivation) in sand-wave deposits, can infer much about the character of ancient tides. In the case of the Folkestone Beds, in which there are cross-bedded sands with mud drapes, the tide seems to have had a strong time–velocity asymmetry, permitting only one episode of sand movement in each tidal cycle, and to have been diurnal with a comparatively small bidiurnal inequality. The apparent spring–neap cycles are of unequal duration, however, suggesting that the lack of some sand layers was due either to the suppression of mud deposition or to a sand transport threshold which for some tides was in excess of the peak speed. The tide had little or no rotary character, a conclusion demanding the restricted seas independently demonstrated for the Folkestone Beds[2,3]. Because time can now be measured from them, we can now be certain that the cross-bedded deposits of the Folkestone Beds represent vastly shorter periods of time than the erosion surfaces which lie between the sets.

I thank J. L. C. McIntyre and R. A. Collins for the opportunity to work at their Pendean Sand Pit.

Received 27 October; accepted 18 November 1980.

1. Casey, R. *Palaeontology* **13,** 487–621 (1961).
2. Narayan, J. *Nature* **199,** 1246–1247 (1963).
3. Narayan, J. *Sediment. Geol.* **6,** 73–109 (1971).
4. Allen, J. R. L. & Narayan, J. *Geologie Mijnb.* **43,** 451–461 (1964).
5. Allen, J. R. L. *Sediment. Geol.* **26,** 281–328 (1980).
6. Allen, J. R. L. *Geol. Mag.* **117,** 437–446 (1980).
7. McCave, I. N. in *Estuarine Hydrography and Sedimentation* 170 (ed. Dyer, K. R.) (Cambridge University Press, 1979).
8. Bagnold, R. A. *The Physics of Blown Sand and Desert Dunes,* 69 (Methuen, London 1941).
9. Gadd, P. E., Lavelle, J. W. & Swift, D. J. P. *J. sedim. Petrol.* **48,** 239–252.
10. Allen, J. R. L. & Friend, P. F. *Sedimentology* **23,** 329–346 (1976).
11. Doodson, A. T. & Warburg, H. D. *Admiralty Manual of Tides,* (HMSO, London, 1941).
12. Scruton, C. T. in *Tidal Friction and the Earth's Rotation,* 154 (eds Brosche, P. and Sundermann, J.) (Springer, Berlin, 1978).
13. Lambeck, K. *The Earth's Variable Rotation: Geophysical Causes and Consequences* (Cambridge University Press, 1980).
14. Ford, J. G. *Mercian Geol.* **5,** 205–240 (1975).

Plutonium transport to and deposition and immobility in Irish Sea intertidal sediments

S. R. Aston* & D. A. Stanners*

Department of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ, UK

The deposition and possible mobility of plutonium in sediments contaminated by nuclear fuel reprocessing wastes are of considerable importance in assessing the short- and long-term fate of this pollutant. Previous studies have shown that marine sediments very effectively and quickly remove plutonium[1], and have suggested both immobile[2] and mobile[2,3] behaviour of plutonium after deposition. Here, the deposition characteristics and lack of vertical migration of plutonium isotopes in intertidal sediments of the Irish Sea, which are contaminated with radioactive wastes from the Windscale reprocessing facility, are discussed.

Plutonium isotopic composition profiles in the accumulating sediments show that plutonium deposition is controlled by the physical deposition of contaminated sediment particles, and the total plutonium activity profiles reflect the general pattern of changes in the quantities of Windscale plutonium discharged. The use of the ^{238}Pu/$^{239+240}$Pu ratios in sediment profiles to estimate sedimentation rates agrees well with the rates measured using independent fission product profiles. These plutonium isotope ratio profiles suggest that plutonium vertical redistribution in the sediments is exceedingly small. When the apparent diffusion rate (which encompasses all diffusive and advective processes) for plutonium is calculated[4,5], a very small apparent diffusion coefficient D must be invoked for the best fitting of previously measured sediment surface ^{238}Pu/$^{239+240}$Pu ratios[6] to those now observed at various depths in the sediments. While changes in the isotopic ratios are well preserved in the sediments after deposition, the peak in the total concentration of plutonium at depth in the dated sediments is found in sediment layers which were deposited ~2–3 yr later than the peak plutonium release from Windscale was discharged. This feature may be a consequence of the time required for the plutonium to travel from the pipeline to the sediments, or of an exchangeable pool or available pool of sediment in the Irish Sea which has a finite clearance rate for the whole Irish Sea basin. The plutonium in these sediments is mostly associated with iron/manganese non-detrital phases and does not undergo diagenetic changes in the oxic sediments.

Sediment cores were collected during mid-1978 from the intertidal reaches of the River Esk, Ravenglass Estuary, eastern Irish Sea, ~10 km south of Windscale[5–7]. Cores were extracted and divided into segments, the plutonium was extracted and radiochemically separated for electrodeposition and α spectroscopy[8]. Fission products in the sediments were determined by γ spectroscopy.

Table 1 shows the ^{238}Pu and $^{239+240}$Pu activities measured in three cores of intertidal sediments together with the profiles of the isotopic ratio ^{238}Pu/$^{239+240}$Pu. Maximum plutonium activities for both ^{238}Pu and $^{239+240}$Pu occur at depths of 3, 3 and 13 cm in cores E-1, E-2 and E-3 respectively. In core E-3 a minor peak in plutonium activities is also found at 36–42 cm. The ratio of ^{238}Pu/$^{239+240}$Pu decreases from surface values of ~0.25 to <0.14 as sediment depth increases. In an earlier investigation of plutonium in sediments from the same field site, Hetherington[1,6] found an approximately exponential increase in plutonium activities towards the surface of the sediment profiles. Hetherington's cores were collected in 1973 and he attributed the profiles to the deposition of contaminated particles which reflected the continuously increasing annual discharge of plutonium from Windscale. The contrasting pattern of plutonium deposition found in the present study is easily accounted for by the decrease in the plutonium discharges from Windscale since 1973.

The total ^{238}Pu and $^{239+240}$Pu activities discharged over the period 1970–78 were 936, 1,128, 1,548, 1,776, 1,248, 1,195, 1,272, 984 and 1,572 Ci yr^{-1} respectively[6,9]. The present core profiles show that the general features of the continuing temporal pattern of plutonium discharge are preserved within the intertidal sediment profiles. The presence of 'buried' plutonium peak activities within the cores confirms that the physical sedimentation of contaminated grains is the overall mechanism of plutonium deposition. Hetherington's data[6], with its increasing plutonium activities towards the sediment surfaces did not definitively show this because the 1973 profiles can be interpreted as being produced by an upward vertical migration of plutonium through the sediment. Plutonium profiles in sediment cores from nearshore, continental shelf and continental slope areas of the eastern US have been attributed to biological mixing and chemical mobilization processes[2,3]. The upward translocation of plutonium in sediments, and the importance of bioturbation on fallout plutonium in slowly depositing nearshore sediments, have been emphasized[3]. However, it has been pointed out that these conclusions cannot be extended to regimes with higher sedimentation rates. The present intertidal areas are ones of high sediment accumulation, and these net rates have been determined as 16, 16 and 65 mm yr^{-1} (±15%) for cores E-1, E-2, E-3 respectively using fission product isotopic ratios[5]. Using a similar technique[5] for the plutonium isotopic ratio in which the sediment profile and previously measured surface sediment ratios (Table 2) are matched by an optimization routine for least squares fitting with an assumed range of net sedimentation rates, the respective rates were 13, 10 and 62 mm yr^{-1}. This shows good agreement with the independent fission product method. An important point in the technique is that the best agreement was achieved when the apparent diffusion coefficient for the plutonium in the sediments was set at $D < 10^{-12}$ cm^2 s^{-1}.

This information is used here to argue that plutonium translocation is very small in these sediments, and the data shown in Table 1 when interpreted in the context of these sedimentation rates indicate that there is a finite time lag of 2–3 yr between the

---

* Present addresses: International Laboratory of Marine Radioactivity, Musée Océanographique, Monaco (S.R.A.); Chemistry Division, JRC, Ispra, Varese 21020, Italy (D.A.S.).

0028–0836/81/070581–02$01.00

Table 1 Activity (pCi g^{-1}) profiles of ^{238}Pu and $^{239+240}Pu$ and the ratio $^{238}Pu/^{239+240}Pu$ in intertidal sediment cores from the River Esk Estuary

Core E-1				Core E-2				Core E-3			
Depth (cm)	^{238}Pu	$^{239+240}Pu$	$^{238}Pu/^{239+240}Pu$	Depth (cm)	^{238}Pu	$^{239+240}Pu$	$^{238}Pu/^{239+240}Pu$	Depth (cm)	^{238}Pu	$^{239+240}Pu$	$^{238}Pu/^{239+240}Pu$
0– 1	21.0±2.4	87.3±6.7	0.24	0– 2	21.7 ±0.9	82.1±2.6	0.27	0– 2	20.9±0.3	79.2±0.7	0.26
1– 2	29.9±3.1	116.7±9.0	0.25	2– 4	23.7 ±1.0	97.0±3.0	0.24	4– 6	16.3±0.2	68.5±0.5	0.24
2– 3	30.5±2.1	124.6±6.3	0.26	4– 6	18.6 ±1.4	89.3±4.5	0.21	8–10	18.2±0.6	76.4±1.6	0.24
3– 4	30.9±2.8	131.5±8.7	0.23	6– 8	11.8 ±0.6	67.7±2.3	0.17	12–14	18.4±0.6	81.7±0.1	0.23
4– 5	26.2±2.3	128.0±8.0	0.20	8–10	2.35±0.17	27.3±0.9	0.09	16–18	11.7±0.3	56.4±0.7	0.21
5– 6	17.8±0.5	96.7±1.6	0.15	10–12	0.94±0.03	17.5±0.3	0.05	20–22	11.0±0.6	42.8±0.6	0.25
6– 7	10.3±0.3	68.5±1.2	0.13	12–14	0.48±0.02	12.4±0.2	0.04	24–26	8.2±0.1	41.3±0.3	0.20
7– 8	7.4±0.5	57.3±2.1	0.12	14–16	0.30±0.03	10.8±0.3	0.03	28–30	2.9±0.1	12.8±0.2	0.22
8– 9	4.6±0.2	38.4±0.9	0.09					32–34	1.5±0.2	8.6±0.5	0.18
9–10	3.1±0.2	32.8±0.8	0.09					36–38	3.4±0.1	20.0±0.4	0.16
10–11	2.4±0.1	27.1±0.5	0.09					40–42	2.7±0.1	20.4±0.4	0.13
11–12	2.3±0.2	25.8±1.2	0.09								
12–13	1.8±0.1	24.1±0.7	0.09								
13–14	1.8±0.1	20.3±0.6	0.09								
14–15	1.5±0.04	16.5±0.2	0.09								

plutonium discharge from Windscale and the arrival of plutonium in the sediments. The data indicate that although the general response to increasing and decreasing discharges is reflected in the overall pattern of deposition, the peak activities are delayed either by a slow transit time from Windscale or because of the clearance time for sediments in the Irish Sea basin. In either situation, the fact that plutonium is very rapidly and effectively removed to sediments[1] is consistent with the slow movement of plutonium in association with bottom sediments rather than with the water masses. Other radioisotope transit or delay times from Windscale to the sampling site have been reported as ~1 month (^{106}Ru, ^{144}Ce)[7], 2–3 months ($^{95}Zr/Nb$)[7] and ~1.5 yr (^{134}Cs, ^{137}Cs)[10]. Clearly the different chemical natures of the radionuclides in the marine environment influence their rates of movement in coastal waters.

Environmental processes are extremely unlikely to differentiate ^{238}Pu and $^{239+240}Pu$, and this allows the isotopic ratio to be used to specify the 'age' of plutonium in a sediment layer. Relatively early discharges of plutonium from Windscale, for example, pre-1968 discharges were significantly smaller than those in more recent years[9,11], and exhibited characteristically low $^{238}Pu/^{239+240}Pu$ ratios. Table 2 shows the $^{238}Pu/^{239+240}Pu$ ratios in surface (<1 cm) sediments from the River Esk estuary as an indication of the changing isotopic composition up to the present higher ratios of ~0.25. The preservation of the older low ratios at depth in the cores with their low plutonium activities (Table 1) further suggests and thus supports the above interpretation that mobilization and vertical mixing with sediments containing greater activities and higher $^{238}Pu/^{239+240}Pu$ ratios is negligible. The intertidal sediments are occasionally bioturbated by populations of *Corophium volutator* (Pallas), but the mixing depth that is the extent to which the organisms may move radionuclides and other substances associated with sediment grains[12,13], seems to be less than the annual increments of sediment accumulation at this location, and does not give rise to any effective vertical movement of radionuclides.

Table 2 The ratio of $^{238}Pu:^{239+240}Pu$ activities in surface (<1 cm) intertidal sediments of the River Esk estuary between 1966 and 1975[6]

Year	$^{238}Pu/^{239+240}Pu$	Year	$^{238}Pu/^{239+240}Pu$
1966	0.053	1972	0.196
1967	0.067	1973	0.200
1968	0.074	1974	No data
1969	0.100	1975	0.238
1970	0.116		
1971	0.185		

Data on the annual $^{238}Pu/^{239+240}Pu$ discharge ratio are not available for this time period.

Theoretical arguments[14] have been used to show that plutonium in marine particulates should be in the Pu (IV) state as insoluble hydroxides, probably in association with iron and manganese oxide/hydroxide colloids. This is strongly supported by the field observations on the Pu (IV) oxidation state in suspended solids in the Irish Sea[15]. A chemical leaching technique[16] was applied to the present sediment cores, and the results showed that 75–100% of the plutonium was in the non-detrital iron/manganese phases. The proportion of plutonium in these phases varies with depth, but there is no apparent trend with depth indicating a lack of diagenetic change in plutonium phase association after deposition. These results agree well with data on fallout plutonium in lacustrine and coastal sediments[17].

From the present results we conclude that: (1) The deposition of contaminated sedimentary particles is the principal control of plutonium input to intertidal sediments in the Windscale area, confirming Hetherington's earlier hypothesis[1,6]. (2) A transit time or clearance rate in the sediment pool of the Irish Sea basin delays the deposition of plutonium at the study site by 2–3 yr. (3) The activities of plutonium and the $^{238}Pu/^{239+240}Pu$ ratio in vertical profiles suggest that post-depositional migration of this pollutant is negligible (apparent $D < 10^{-12}$ cm^2 s^{-1}). (4) The plutonium is mainly in the non-detrital iron/manganese phase, and there is no evidence for diagenetic changes in phase association after burial in the oxic sediment layer.

We thank the NERC for financial assistance, the Fisheries Radiobiological Laboratory, MAFF (Lowestoft) for use of facilities, and M. Kelly and D. J. Assinder for helpful comments.

Received 13 August; accepted 27 November 1980.

1. Hetherington, J. A. in *Environmental Toxicity of Aquatic Radionuclides: Models and Mechanisms* 81–106 (Ann Arbor, Michigan, 1976).
2. Livingston, H. D. & Bowen, V. T. in *Environmental Toxicity of Aquatic Radionuclides: Models and Mechanisms*, 107–130 (Ann Arbor, Michigan, 1976).
3. Livingston, H. D. & Bowen, V. T. *Earth planet. Sci. Lett.* **43,** 29–45 (1979).
4. Lerman, A. in *Proc. 3rd natn. Symp. Radioecol. NTIS*, Springfield, Virginia, 936–944 (1971).
5. Aston, S. R. & Stanners, D. A. *Estuar. Coastal mar. Sci.* **9,** 529–541 (1979).
6. Hetherington, J. A. *Mar. Sci. Commun.* **4,** 239–274 (1978).
7. Hetherington, J. A. & Jefferies, D. F. *Neth. J. Sea Res.* **8,** 319–338 (1974).
8. Talvitie, N. A. *Analyt. Chem.* **44,** 280–283 (1972).
9. British Nuclear Fuels Ltd, *Annual Report on Radioactive Discharges and Monitoring of the Environment* (BNFL, Risley, 1978, 1979).
10. Stanners, D. A. & Aston, S. R. *Estuar. Coastal mar. Sci.* (in the press).
11. Hetherington, J. A., Jefferies, D. F., Mitchell, N. T., Pentreath, R. J. & Woodhead, D. S. in *Transuranium Nuclides in the Marine Environment*, 139–153 (IAEA, Vienna, 1976).
12. Robbins, J. A., Krezoski, J. R. & Mozley, S. C. *Earth planet. Sci. Lett.* **36,** 325–333 (1977).
13. Aller, R. C. thesis, Yale Univ. (1977).
14. Aston, S. R. *Mar. Chem.* **8,** 319–325 (1980).
15. Nelson, D. M. & Lovett, M. B. *Nature* **276,** 599–601 (1978).
16. Chester, R. & Hughes, M. J. *Chem. Geol.* **2,** 249–262 (1967).
17. Alberts, J. J., Muller, R. N. & Orlandini, J. A. in *Radiological and Environmental Research Division A. Rep.* ANL-76-88 Part III (Argonne National Laboratory, 1976).

Function of the mandibular tooth comb in living and extinct mammals

Kenneth D. Rose & Alan Walker

Department of Cell Biology and Anatomy, The Johns Hopkins University, School of Medicine, Baltimore, Maryland 21205

Louis L. Jacobs

Museum of Northern Arizona, Flagstaff, Arizona 86001

Among the most interesting mammalian dental specializations is the mandibular 'tooth comb' or 'tooth scraper' that evolved independently in certain primates and other mammals. Its occurrence is most widely known in lemurs and lorises, where it is comprised of the long, slender, procumbent incisors (one or two pairs) and incisiform canines (Fig. 1). In non-primates the canines are not incorporated into the comb. Some tree shrews (Tupaiidae) possess a tooth comb consisting of the four central incisors, and some early Tertiary arctocyonid condylarths had a similar structure composed of all six lower incisors[1]. The extant flying lemurs (Dermoptera: *Cynocephalus*) also have a dental 'comb' but it is very different from the ones already mentioned, consisting of two pairs of pectinate incisors, each tooth modified into a comb with as many as 15 tines. This condition, although sometimes said to be similar to that in lemurs, is unique to *Cynocephalus*. One of the principal functions of the tooth comb in primates is to comb the fur, and we present here indirect evidence that condylarths used this structure in the same way, millions of years before tooth combs evolved in prosimians. We also show that the comb-like incisors of *Cynocephalus*, contrary to popular belief, probably do not function to comb the fur.

The function of the mammalian tooth comb has been controversial. Its comb-like appearance in prosimian primates led very early to the suggestion that it is used in grooming the fur[2], a hypothesis that persisted unchallenged for more than a century[3–6]. Some authors have doubted[7] or even denied[8] such a function for the tooth comb, but behavioural studies have confirmed its use during grooming[9–12]. Tupaiids have been reported to use their tooth combs for grooming in the same manner as do lemurs[6]. It is now recognized that grooming is not the sole function of the prosimian tooth comb. Several smaller lemurs and lorises (for example, *Phaner furcifer, Euoticus elegantulus* and *Galago senegalensis*) use this structure in procuring gum[13–15], and the sifaka, *Propithecus verrauxi*, uses its comb to gouge bark and dead wood, particularly in the dry season[16,17]. Some lemurs (such as *Microcebus murinus, Propithecus verrauxi* and *Indri indri*) scoop fruit pulp with their dental combs[12,18].

There has been disagreement as to whether prosimians, when grooming, use their tooth combs to comb the fur or merely to scrape it. Buettner-Janusch and Andrew[12] found fine hairs caught between the incisors of *Galago* after grooming—compelling indirect evidence of combing. One of us (A.W.) has observed hair-combing with the tooth comb in several captive lemurs and lorises. Martin[18], citing observations of R. Every, reported that teeth in the comb of lemurs bear " 'whip marks' produced by hairs pulled between the teeth".

Using scanning electron microscopy (SEM) to examine teeth from prosimian combs, we have confirmed the presence of fine vertical grooves (~10–20 μm wide, also visible with the light microscope), presumably made by hairs, on the lateral and medial sides of the teeth and on their median dorsal (lingual) ridges (Figs 1, 2). The grooves are discontinuous, being absent from the intervening concave surfaces between the sides of each tooth and the median ridge. Within these grooves are long, finer microstriations (<1 μm wide), which may have resulted from scratching by the cuticular layer of the hair. Similar grooves are visible on the incisors of some tupaiids (for example, *Tupaia glis, Tupaia tana*).

When did the prosimian tooth comb evolve, and for how long has it been involved in grooming? There is no tooth comb in early Tertiary Adapidae (considered the oldest lemuriform primates by many authors), although one species, *Adapis parisiensis*, shows modifications of the lower anterior teeth that may foreshadow a tooth comb[19]. Aside from adapids, the oldest known lemuriforms are early Miocene lorisoids from East Africa[20,21] and late Miocene lorisoids from Asia[22]. Crowns of the anterior teeth are not yet known for early Miocene lorisoids, but the configuration of canine and incisor roots in some specimens indicates that the tooth comb was fully developed by that time[23]. The most ancient lorisoid in which the crowns of the comb teeth are known (and also the only fossil one in which they are known) is *Nycticeboides simpsoni* (holotype, YGSP 8091) from the Siwalik Group, Pakistan, probably between 7 and 8 million years old[22,24]. These comb teeth, including the crowns of

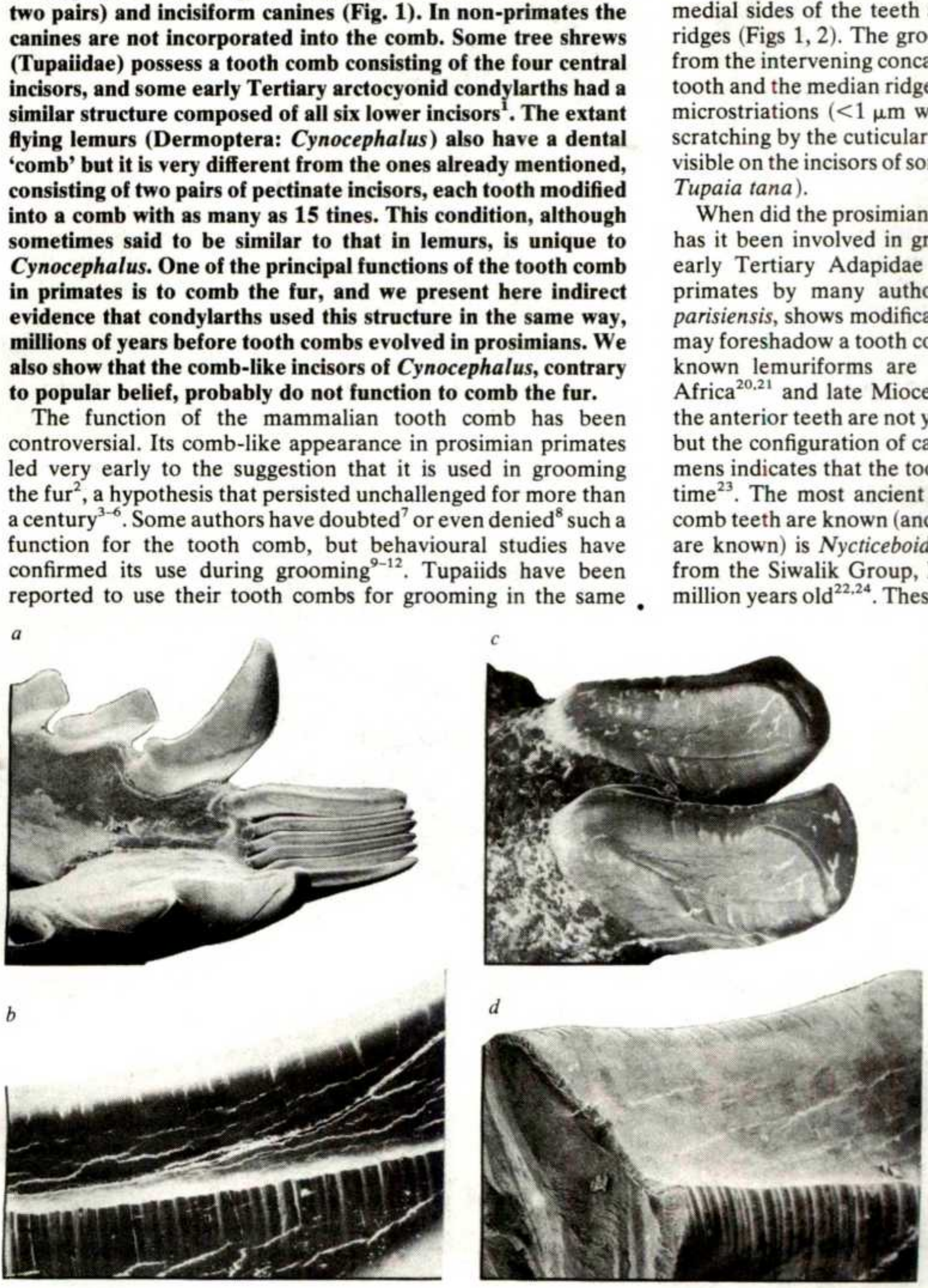


Fig. 1 *a*, Mandibular comb of *Euoticus elegantulus*. *b*, Left central incisor of *Galago crassicaudatus*, showing hair grooves on the median ridge (above) and on the interstitial facet (below). *c*, Left incisors of a middle Palaeocene arctocyonid from Montana (Carter County Museum no. 79-5) showing hair grooves on interstitial facet and extending across extensive dentine exposure on median ridge. *d*, Left incisor of an early Eocene arctocyonid (cf. *Thryptacodon antiquus*, Princeton University no. 20853) from Wyoming, showing hair grooves almost identical to those in the modern *Galago*.

0028–0836/81/070583–03$01.00

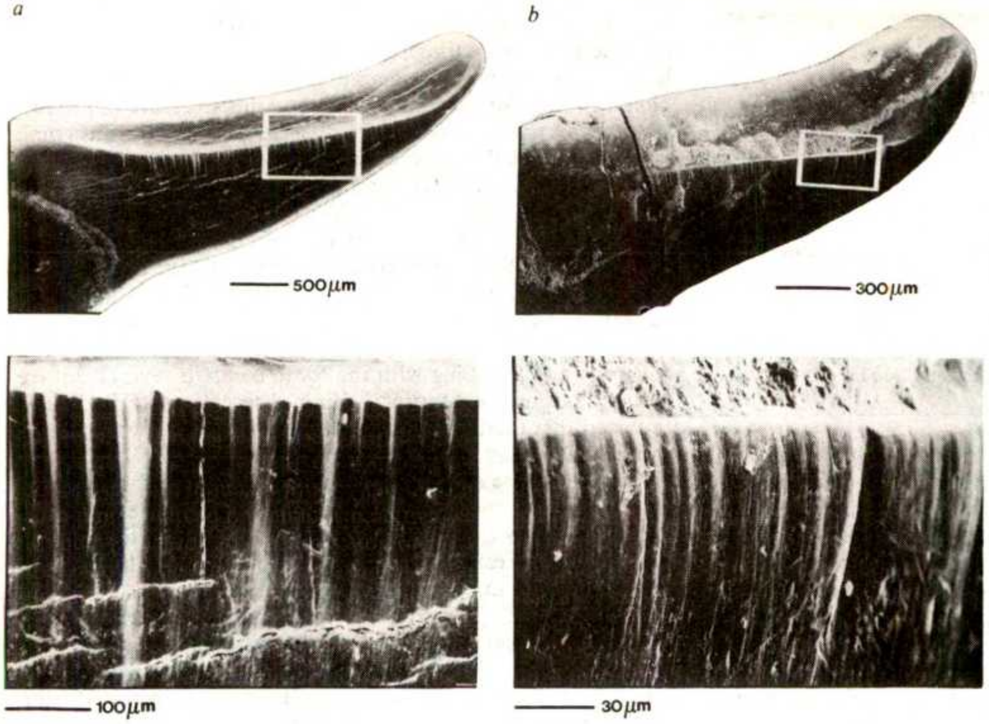


Fig. 2 *a*, Left incisor of *Galago crassicaudatus* showing hair grooves on the interstitial facet. *b*, Comparative scanning electron micrographs of the left incisor of the type of *Nycticeboides simpsoni*, showing similar grooves.

one canine and two incisors, are morphologically very similar to corresponding teeth in modern lorises and galagos. Scanning electron micrographs of these comb teeth reveal grooves (10–15 μm wide) and finer scratches (<1 μm) remarkably similar to those formed on the comb teeth of modern lorisoids that use the tooth comb for grooming (Fig. 2). Thus it seems that grooming behaviour like that performed by extant lorisoids had evolved by the late Miocene. The type specimen of *N. simpsoni* documents, therefore, the earliest record of tooth comb grooming behaviour in prosimian primates.

Gingerich and Rose[1] reported the occurrence of a lemur-like tooth comb in an earliest Eocene (Clarkforkian) arctocyonid condylarth (cf. *Thryptacodon* from Wyoming) and postulated that it functioned in grooming. SEM examination of this specimen reveals the presence of closely spaced fine grooves (~25 μm wide), virtually identical in size and form to those on prosimian comb teeth (Fig. 1). These grooves indicate that *Thryptacodon*, like lemurs, combed its fur with its tooth comb.

A recently discovered specimen of a middle Palaeocene (Torrejonian) arctocyonid from Montana possesses a tooth comb like that in the Eocene specimen, but with larger incisors that are more heavily worn. SEM examination of this specimen also reveals grooves on the sides and median dorsal ridge of the incisors (Fig. 1). That these grooves formed during life is demonstrated by their continuation across the enamel edges and dentine exposure of the median dorsal ridge of each incisor. The

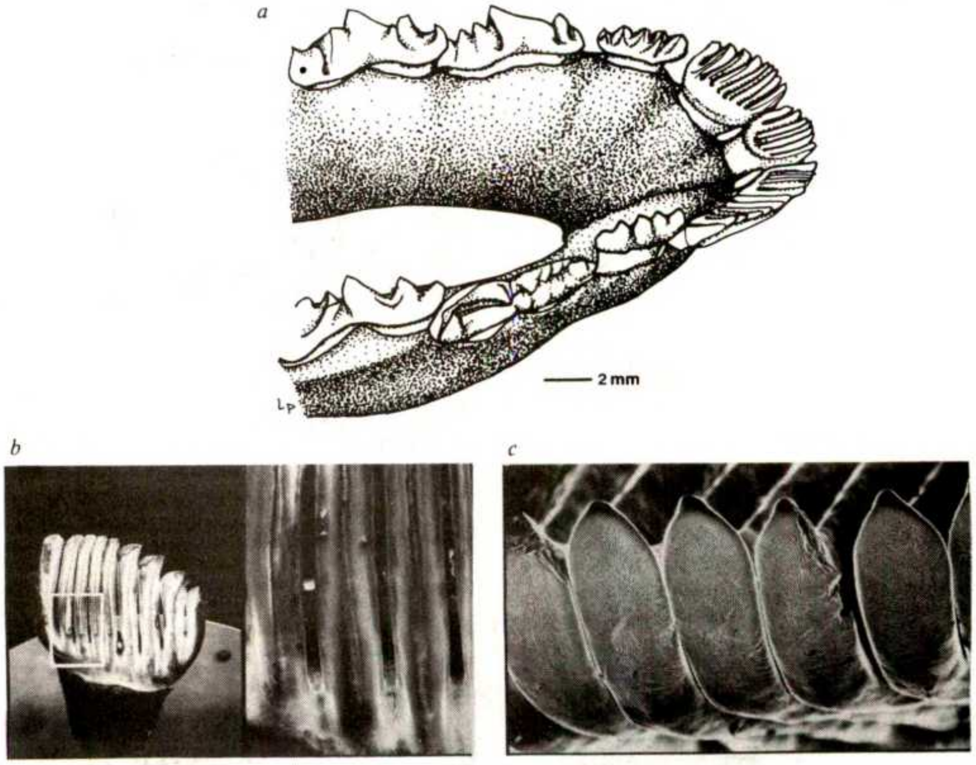


Fig. 3 *a*, Drawing of the anterior dentition of *Cynocephalus variegatus*, showing comb-like incisors. *b*, Scanning electron micrographs of a central incisor of *Cynocephalus* (left). Note that no grooves are evident between the tines of the tooth. *c*, Scanning electron micrograph of labial apical facet on incisor of *Cynocephalus*, showing exposure of dentine tubules.

grooves on these teeth are considerably broader (~100 µm wide) than those on the incisors of cf. *Thryptacodon*, and probably represent the coalescence of several smaller grooves—a consequence of much heavier wear in this specimen.

Early Eocene (Wasatchian) specimens of the arctocyonid *Chriacus* from Wyoming also possess an incisor comb very similar to that in cf. *Thryptacodon* and, similarly, minute grooves can be detected on the sides of the incisors[25].

These arctocyonid specimens document the earliest known occurrence in the mammalian fossil record of grooming with a dental comb, predating the oldest evidence of a tooth comb in fossil primates by at least 40 Myr. They may represent three different genera of smaller arctocyonids, suggesting that fur-combing with the incisors was an important behavioural trait in several early Tertiary condylarths.

Recently primatologists have concentrated on the question of the original function of the tooth comb. Martin[18,26] proposed that the tooth comb initially functioned in procuring gum, especially in smaller species, and that grooming was a secondary function. Other authors have contended that the comb must have evolved primarily as a grooming apparatus[12,27]. The arctocyonid tooth combs described above are relevant to this problem. Grooming was clearly performed by even the oldest mammal (middle Palaeocene) known to possess a tooth comb, but this form also shows heavy wear of the incisors, especially at the apices. The tips of the incisors seem to have formed a single cutting edge, similar to the condition in bark-gouging primates such as *Propithecus*. Thus the oldest available evidence indicates that the tooth comb functioned for both grooming and food procurement. The specimen of cf. *Thryptacodon* is less worn than the Palaeocene specimen, and its incisors are relatively longer and more slender. Moreover, they are pointed at the tips, as in *Lemur*, and do not form a single cutting edge. Eocene *Chriacus* resembles cf. *Thryptacodon* in these features. Thus, in cf. *Thryptacodon* and Eocene *Chriacus*, food procurement may have become subordinate to grooming, or may no longer have been a significant function of the tooth comb. The question of whether gum-eating or grooming was the earliest function of the combs of the ancestors of lemurs and lorises is still unanswered, as even those species that eat gum today seem to have no distinctive wear resulting from that activity.

The function of the comb-like incisors of *Cynocephalus* (Fig. 3) is unclear, but it has been widely stated that they, too, are used to comb the fur[28-30]. A few authors disagree with this view[8] or consider that grooming is not the principal function[31,32], and we know of no documented observations of flying lemurs using their incisors for combing. Indeed, Wharton[33] reported that *Cynocephalus* grooms by licking its fur; he did not mention use of the incisors—the incisors probably serve in food procurement. *Cynocephalus* is strictly herbivorous, feeding largely on leaves, blossoms and fruits[30,33,34], and it has even been suggested that the incisors are used for 'scraping the green colouring matter out of leaves'[35]. Though this is difficult to substantiate, stomach contents we examined were dominated by small bits of leafy matter, confirming Wharton's[33] observations. The lower incisors may function to crop leaves against the edentulous part of the premaxilla (as in many artiodactyls)[36], or to scrape leaves[29].

Our SEM examination of incisors of both extant species of *Cynocephalus* (*C. volans* and *C. variegatus*) revealed no evidence of the fine grooves seen on incisors of lemurs, lorises, tupaiids and arctocyonids, and inferred to be caused by hair-combing. A wear facet is present on most specimens on the labial (ventral) side of the tooth, at the apices of the tines. It probably results from contact with the edentulous part of the premaxilla and is consistent with a cropping function for the dermopteran tooth comb. The precise function of the elaborate tooth comb of *Cynocephalus* remains uncertain, but it seems very unlikely that this apparatus, the most comb-like of such structures, is actually used to comb the fur.

We thank D. Baird, P. Gingerich, M. Lambert, D. Pilbeam, R. Thorington, M. Turner, and the Director, Geological Survey of Pakistan, for access to specimens used in this study, and Linda Perez and Gary Morgan for technical assistance. This work was supported by NSF grant BNS 78-24499 to A.W.

Received 30 July; accepted 18 December 1980.

1. Gingerich, P. D. & Rose, K. D. *J. Mammal.* **60,** 16–22 (1979).
2. Cuvier, F. & St Hilaire, G. *Histoire Naturelle des Mammifères* (Bull. Mus. nat. Hist., Paris, 1825).
3. Jones, F. Wood *J. Anat.* **52,** 345–353 (1918).
4. Jones, F. Wood *Proc. R. Soc. Tasmania 1925*, 29–30 (1925); *Man's Place Among the Mammals* (Longmans, London, 1929).
5. Clark, W. E. Le Gros *Early Forerunners of Man* (William Wood, Baltimore, 1934).
6. Clark, W. E. Le Gros *The Antecedents of Man* (Edinburgh University Press, Edinburgh, 1959).
7. Avis, V. *Am. J. phys. Anthrop.* **19,** 55–61 (1961).
8. Stein, R. M. *Am. Nat.* **70,** 19–28 (1936).
9. Lowther, F. de L. *Zoologica* **24,** 477–480 (1939).
10. Roberts, D. *J. Anat.* **75,** 236–238 (1941).
11. Doyle, G. A. in *Prosimian Biology* (eds Martin, R. D., Doyle, G. A. & Walker, A. C.) 213–231 (Duckworth, London, 1974).
12. Buettner-Janusch, J. & Andrew, R. J. *Am. J. phys. Anthrop.* **20,** 127–129 (1962).
13. Petter, J.-J., Schilling, A. & Pariente, G. *Terre Vie* **25,** 287–327 (1971).
14. Charles-Dominique, P. *Biol. Gabonica* **7,** 121–228 (1971); in *Prosimian Biology* (eds Martin, R. D., Doyle, G. A. & Walker, A. C.) 131–150 (Duckworth, London, 1974); *Ecology and Behavior of Nocturnal Prosimians* (Duckworth, London, 1977).
15. Hladik, C. M. in *The Study of Prosimian Behavior* (eds Doyle, G. A. & Martin, R. D.) 307–357 (Academic, New York, 1979).
16. Petter, J.-J. *Mem. Mus. nat. Hist. Nat.* Ser. A, **27,** 1–146 (1962).
17. Richard, A. *Folia primatol.* **22,** 178–207 (1974); in *The Ecology of Arboreal Folivores* (ed. Montgomery, G.) 519–533 (Smithsonian Inst. Press, Washington, 1978).
18. Martin, R. D. in *The Study of Prosimian Behavior* (eds Doyle, G. A. & Martin, R. D.) 45–77 (Academic, New York, 1979).
19. Gingerich, P. D. in *Lemur Biology* (eds Tattersall, I. & Sussman, R.) 65–80 (Plenum, New York, 1975).
20. Simpson, G. G. *Bull. Mus. comp. Zool.* **136,** 39–62 (1967).
21. Walker, A. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 90–99 (Harvard University Press, Cambridge, Massachusetts, 1978).
22. Jacobs, L. L. *Nature* **289,** 585–587 (1981).
23. Walker, A. *Uganda J.* **33,** 90–91 (1969).
24. Tauxe, L. *Nature* **282,** 399–401 (1979).
25. Rose, K. D. & Bown, T. M. (in preparation).
26. Martin, R. D. *Phil. Trans. R. Soc.* B**264,** 295–352 (1972).
27. Szalay, F. S. & Seligsohn, D. *Folia primatol.* **27,** 75–82 (1977).
28. Peyer, B. *Comparative Odontology* (University Chicago Press, Chicago, 1968).
29. Vaughan, T. *Mammalogy* (Saunders, Philadelphia, 1978).
30. Walker, E. P. *et al. Mammals of the World* (Johns Hopkins University Press, Baltimore, 1975).
31. Jepsen, G. L. *Peabody Mus. nat. Hist. A. Rep.* **3,** 1–14 (1962).
32. Chapman, H. C. *Proc. Acad. nat. Sci.* **54,** 241–254 (1902).
33. Wharton, C. H. *J. Mammal.* **31,** 269–273 (1950).
34. Lim Boo Liat *Int. Zoo Yb.* **7,** 196–197 (1967).
35. Raven, H. C., cited in Gregory, W. K. *Evolution Emerging* 387 (Macmillan, New York, 1951).
36. Winge, H. *The Interrelationships of the Mammalian Genera* Vol. 1 (Reitzels, Copenhagen, 1941).

Miocene lorisid primates from the Pakistan Siwaliks

Louis L. Jacobs

Department of Geosciences, University of Arizona, Tucson,
Arizona 85721 and Museum of Northern Arizona,
Flagstaff, Arizona 86001

In 1975–79, joint expeditions from Yale University and the Geological Survey of Pakistan (YGSP) recovered fossil lorisids in the Siwalik Group of Pakistan from four localities, spanning a period before 10 Myr ago to about 7 Myr ago[1,2]. In three of the localities, only isolated teeth or fragments were found, whereas the fourth and youngest locality yielded dental, cranial and some postcranial remains of a single individual described here as a new genus and species. These specimens are the first fossil lorisids known from outside East Africa, and include the only recovered postcranial remains from slow-moving arboreal lorisines. The findings indicate that significant tracts of forests in the Siwalik environmental mosaic may have been utilized by hominoid primates, notably *Ramapithecus*.

Modern lorisids (the galagos, bushbabies, pottos and lorises) occur today in Asia, from forested India eastwards, and in Africa. Two genera of slow-moving lorises occur in Asia, and about four genera in Africa, two of which are slow movers and two others leapers. The slow movers (or lorisines) and the leapers (or galagines) can be distinguished by dental and especially postcranial morphology[3]. Until recently, the fossil record

0028-0836/81/070585-03$01.00

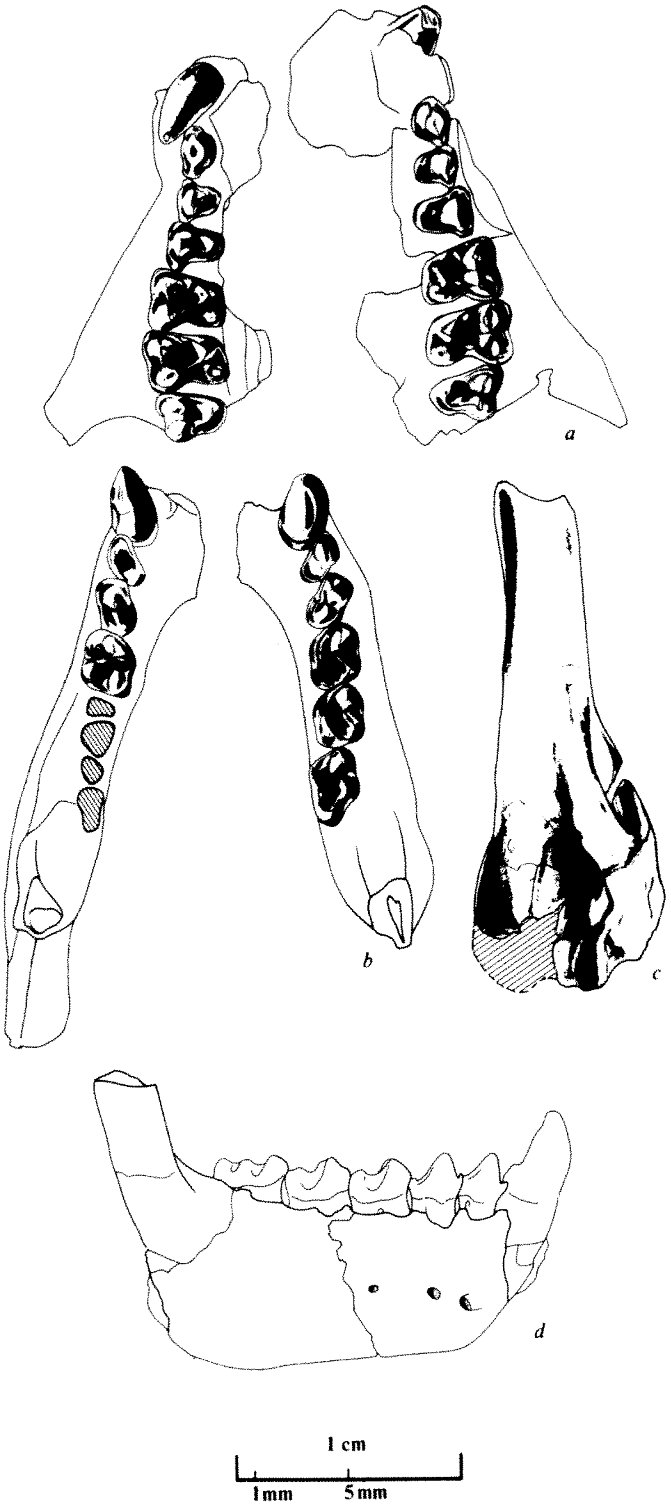


Fig. 1 *Nycticeboides simpsoni*, new genus and species, from the Pakistan Siwaliks: *a*, upper dentition; *b*, lower dentition; *c*, distal humerus; *d*, right dentary in labial view.

of lorisids was confined to East Africa, in deposits of Miocene–Pleistocene age[3–5]. The oldest Asian find, a fragmentary tooth (YGSP 8094), is from YGSP locality 259 in the Siwaliks[6,7]. (*Indraloris* was described as a Siwalik lorisid by Lewis. Known specimens, including many that are undescribed, do not have any features that place them in the Lorisidae; rather, Gingerich and Sahni have placed *Indraloris* with *Sivaladapis* in the Adapidae.) Two other localities (YGSP localities 450 and 182A) have produced isolated lorisid teeth. Hominoid primates, including *Ramapithecus*, and tree shrews are found at or near these three localities[8]. The ages of the localities are between 10 and 8 Myr ago, based on associated faunal and palaeomagnetic data[1,2]. Tree shrews now live in forested regions of Asia from India eastwards.

The youngest Siwalik locality to produce a lorisid is YGSP locality 363. It occurs higher in the section than the last occurrence of hominoids, and therefore is probably younger than 8 Myr (ref. 2). The lorisid remains are from one individual, and no other taxa have been recovered from the locality. It represents a new genus and species *Nycticeboides simpsoni.* This taxon (Fig. 1) is known only from the holotype (YGSP 8091), which includes most of the dentition with the exception of upper incisors and parts of the anterior lower dentition. Some skull fragments and a few postcranial fragments, mostly unprepared, are also represented. *N. simpsoni* is a lorisid because it has a tooth comb[9], the mandibular symphysis is not fused, the morphology of the cheek teeth is similar to and the dental formula apparently the same as those of other lorisids (see ref. 10). It is referred to the subfamily Lorisinae because the premolars are relatively simple, the postorbital bar is strong, the mandible heavy, and the humerus has a weak ectepicondylar flange. The generic name reflects similarities to *Nycticebus* and the species is named after Dr George Gaylord Simpson.

For the present, the diagnosis is the same for both genus and species because this is a monotypic taxon. The dentary is heavy with a vertical, robust ascending ramus where it joins the horizontal ramus. The cheek teeth are relatively simple, without strong cingula. The premolars are not strongly molariform. Among living lorises, *Nycticeboides* is most similar to *Nycticebus* in the shape of the base of the ascending ramus and in the morphology of P_3 and P_4. However, *Nycticeboides* has weaker labial cingula on molars and stronger styles on P^3 and P^4. It is similar in size to *Nycticebus coucang pygmaeus* (Table 1; for measurements of *N. coucang pygmaeus*, see ref. 11). *Nycticeboides* is distinct from *Loris* in having a narrow P^4, more rounded cusps on the cheek teeth and better developed cingula on P_3 and P_4. The hypocone on upper molars is stronger and the P^3 more molariform in *Nycticeboides* than *Perodicticus. Arctocebus* and *Mioeuoticus* are distinct from *Nycticeboides* in having

Table 1 Measurements of *Nycticeboides simpsoni*, new genus and species

	Length (mm)	Width (mm)
Right P^2	2.2	1.6 (estimated)
Left P^2	2.2	1.7
Right P^3	1.9	1.9
Left P^3	2.1	1.9
Right P^4	2.1	2.5
Left P^4	2.1	2.5
Right M^1	2.7	3.4
Left M^1	2.7	3.4
Right M^2	2.7	3.9
Left M^2	2.7	3.7
Right M^3	2.2	3.1
Left M^3	2.3	3.0
Right P_2	2.6	2.0
Left P_2	2.7	2.1
Right P_3	2.0	1.9
Left P_3	2.1	2.0
Right P_4	2.2	2.0
Left P_4	2.4	2.0
Right M_1	2.7	2.2 trigonid 2.3 talonid
Left M_1	2.7	2.2 trigonid 2.3 talonid
Right M_2	2.9	2.3 trigonid 2.2 talonid
Right M_3	3.5	2.2 trigonid 2.0 talonid
Depth of ramus below right P_4	6.0	
Depth of ramus below right M_1	5.9	
Depth of ramus below right M_2	6.3	
Depth of ramus below right M_3	6.9	

more nearly square molars with stronger cingula and, in the case of *Arctocebus*, stronger ridges connecting cusps on molars.

The way in which lorisines and galagines move is reflected in the shape of their humeri. Galagines are vertical clingers and leapers with a fairly broad habitat tolerance[3]. Lorisines are slow-moving arboreal quadrupeds, limited in distribution to fairly dense jungle. The humerus of *Nycticeboides* is more similar to that of modern, slow-moving lorises in having a weak ectepicondylar flange (Fig. 1*c*). In addition, the relative weakness of the lesser tuberosity and the buttress for insertion of the *teres major* resembles lorisines rather than galagines, although this conclusion is less secure because of the lateral distortion of the proximal end of the humerus. However, this distortion would tend to exaggerate similarities with galagos rather than lorises.

N. simpsoni represents a lorisine of modern grade in locomotor adaptation. As a lorisine, it suggests the presence of tracts of jungle large enough to support the existence of slow-moving arboreal quadrupeds in that area of South Asia up to at least 7 Myr ago, regardless of what other components of environmental mosaic were available. If the older lorisids of the Siwaliks were similarly adapted, and if the joint occurrence of taxa at a fossil locality reflects community composition, then the hominoid *Ramapithecus* was probably at least an occasional visitor to forest tracts inhabited by lorises and tree shrews.

Received 4 July; accepted 14 November 1980.

1. Pilbeam, D. *et al. Postilla* **179**, 1–45 (1979).
2. Tauxe, L. *Nature* **282**, 399–401 (1979).
3. Walker, A. in *Primate Locomotion* (ed. Jenkins, F. A.) 349–381 (Academic, New York, 1977); *Am. J. phys. Anthrop.* **33**, 249–262 (1970); in *The Study of Prosimian Behavior* (eds Doyle, G. A. & Martin, R. D.) 543–565 (Academic, New York, 1979).
4. Walker, A. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 90–99 (Harvard University Press, Cambridge, Massachusetts, 1978); in *Prosimian Biology* (eds Martin, R. D., Doyle, G. A. & Walker, A. C.) 435–447 (Duckworth, London, 1974).
5. Simpson, G. G. in *Olduvai Gorge 1951–1961* (ed. Leakey, L. S. B.) 15–16 (Cambridge University Press, 1965).
6. Lewis, G. E. *Am. J. Sci.* **26**, 134–138 (1933).
7. Gingerich, P. D. & Sahni, A. *Nature* **279**, 415–416 (1979).
8. Jacobs, L. L. in *Comparative Biology and Evolutionary Relationships of Tree Shrews* (ed. Luckett, W. P.) 205–216 (Plenum, New York, 1980).
9. Rose, K. D., Walker, A. C. & Jacobs, L. L. *Nature* **289**, 583–585 (1981).
10. Simpson, G. G. *Bull. Mus. comp. Zool.* **136**, 39–62 (1967).
11. Elliot, D. G. *A Review of the Primates* (American Museum of Natural History, 1913).

Immediate phytochrome action in inducting α-galactosidase in lettuce seeds

David W. M. Leung & J. Derek Bewley

Department of Biology, University of Calgary, Calgary, Alberta,
Canada T2N 1N4

The light-induced breaking of dormancy of Grand Rapids lettuce seeds imbibed in darkness at 25 °C is a classical phytochrome-controlled phenomenon[1]. Although it has been shown that irradiation of seeds 1–2 h after the start of imbibition produces the active form of phytochrome (Pfr) which acts rapidly after its formation[2,3], the completion of germination (observed as radicle protrusion) is only evident several hours later. Clearly, then, radicle emergence is temporally removed from the rapid (and presumably primary) action of phytochrome. Very little is known about the processes set in motion by Pfr that lead to the completion of germination. Here we report that red light-induced stimulation of α-galactosidase (EC 3.2.1.22) activity occurs substantially before germination is completed, and show that the action of phytochrome in promoting an increase in the activity of this enzyme is very rapid.

Table 1 Effects of brief red-light irradiation on α-galactosidase activity and germination in lettuce seeds

Treatment	Time after start of imbibition (h)	Enzyme activity (units × 10^4 per seed)	% Germination
Dry seeds	0	12.6	—
Irradiated seeds	5	18.5	0
	12	18.6	10 (100)
Dark-imbibed seeds	5	12.5	0
	12	12.8	2 (8)

Batches of 25–50 lettuce seeds (*Lactuca sativa* cv. Grand Rapids, Ferry Morse Seed Co.) were allowed to imbibe in darkness at 25 °C. After 1.5 h from the start of imbibition some seeds were induced to germinate by 3 min of red light obtained from two 20-W Cool White fluorescent tubes filtered through a Ruby Red Cinemoid filter transmitting maximally at 660 nm (ref. 7). After 5 or 12 h triplicate batches of seeds were collected, their germination scored and they were ground in 0.05 M potassium phosphate buffer (*p*H 6.85) to obtain a crude homogenate. After clearing at 12,000*g* for 20 min at 2 °C the extract was assayed for α-galactosidase activity. The assay mixture contained 0.1 ml each of McIlvaine buffer, *p*H 5.2, *p*-nitrophenyl α-D-galactopyranoside (10^{-3} M) (Sigma) dissolved in buffer, and properly diluted enzyme extract. After incubation at 30 °C for 5–15 min, 1.0 ml of cold Na_2CO_3 (10^{-1} M) was added to terminate the enzyme reaction. Control experiments consisted of adding substrate at the end of the incubation. Enzyme activity was measured by a change in absorbance at 400 nm due to enzymatic release of *p*-nitrophenol. The molar extinction coefficient for *p*-nitrophenol was taken as 1.84×10^4 and 1 unit of enzyme activity is defined as that which brings about the release of 1 μmol of *p*-nitrophenol per min in the above conditions of assay. Variability in enzyme activity between triplicate samples of seeds was within ±0.4 units × 10^4 seed. Values in parentheses show the percentage of germination 24 h after the start of imbibition.

Table 1 shows that lettuce seeds in darkness at 25 °C (that is, dormant seeds) maintain the same level of α-galactosidase as dry seeds, whereas red light promotes both germination and an appreciable increase in α-galactosidase activity. This increase occurs within 5 h of the start of imbibition and within 3.5 h of irradiation. This is well before radicle emergence, which is noticeable some 5–7 h later. Enzyme activity and germination are induced by a 2-min exposure to red light, but not by shorter irradiations (Table 2). The promotional effect of 2-min red light on germination is abolished if far-red light is then applied immediately, but the red light-induced increase in α-galactosidase activity is unaffected.

Table 2 Effects of a single brief far-red light irradiation on enzyme activity in lettuce seeds treated with a single brief red-light irradiation or in seeds imbibed in darkness

Irradiation treatment	% Germination	Enzyme activity (units × 10^4 per seed)
None	8	12.4
Red 1 min	8	12.5
Red 2 min	84	16.8
Red 2 min + far-red 8 min or 20 min	4	16.9
Far-red 8 min or 20 min	5	12.4

Seeds were imbibed in darkness, or given red light 1.5 h after imbibition. One batch of red light-treated seeds was irradiated immediately with either 8 min or 20 min far-red light, along with one batch of dark-imbibed seeds. Seeds were collected at 24 h for counting the percentage germinated, and at 5 h for α-galactosidase determination, as described in Table 1 legend. Far-red irradiation was obtained from a 100-W tungsten filament lamp filtered through Ilford filters 608 and 806 transmitting between 700 and 1,000 nm.

In view of the rapid action of Pfr in lettuce seeds[2,3], and the observation that there is a rapid escape from photoreversibility

0028-0836/81/070587-02$01.00

Table 3 Effects of intermittent brief red light irradiations separated by periods of brief far-red light irradiations or by periods of darkness on lettuce seed germination and levels of α-galactosidase activity

Irradiation treatment	Total equivalent time of red light irradiation (min)	% Germination	Enzyme activity (units × 10^4 per seed)
a, Red 5 s + far-red 25 s, 24×	2	4	16.6
Red 5 s + darkness 25 s, 24×	2	84	16.9
b, Red 2 s + far-red 8 s), 60×	2	4	12.8
Red 2 s + darkness 8 s, 60×	2	80	16.7
c, Red 2 s + far-red 8 s, 30×	1	4	12.6
Red 2 s + darkness 8 s, 30×	1	8	12.6

Seeds imbibed in darkness for 1.5 h were irradiated as shown, and each irradiation sequence was repeated the number of times indicated. The timing of irradiations was within ±0.5 s.

(of the order of seconds) of the potentiation of flowering in *Pharbitis nil*[4], and perhaps of other phenomena[5], we studied the possibility that a rapid Pfr action is also involved in the promotion of α-galactosidase activity. The 2-min promotional period of red light was therefore applied as repeated 2- or 5-s pulses, with intervening longer periods of far-red light or darkness. Two-second pulses of red light (cumulative to 2 min) with intervening dark periods resulted in a high percentage of germination and full promotion of enzyme activity (Table 3*b*). With intervening periods of far-red light, however, both germination and enzyme activity were suppressed. Five-second pulses of red light with intervening dark periods also resulted in high germination and full enzyme activity, but the latter was not suppressed by intervening periods of far-red light (Table 3*a*). If the cumulative time of red light pulses was less than 2 min there was no promotion of germination or enzyme activity (Table 3*c*).

Several important points arise from these experiments. First, red light promotes an increase in the activity of a hydrolytic enzyme, α-galactosidase, before the completion of germination. Second, the promotional effects of red light on germination and α-galactosidase activity are cumulative, provided that the critical total exposure time of 2 min is achieved. However, far-red light can inhibit germination at times when enzyme activity is unaffected, indicating two distinct phytochrome-controlled reactions. Third, the effects of Pfr on the induction of enzyme activity are very rapid, and exposure to red light for only 5 s precludes the possibility that far-red light has any subsequent inhibitory effects, even though it still has a strong negative effect on germination. Thus, the consequences of Pfr action for between 2 and 5 s are irreversible as far as the induction of α-galactosidase activity is concerned. Such a rapid action of phytochrome on enzyme induction is unknown elsewhere, and is unlikely to be explained in terms of direct gene activation[6]. It is not known whether α-galactosidase is involved in dormancy-breaking in lettuce seeds, and its role may be associated with the early mobilization of raffinose reserves[7,8]. Certainly, its presence in the lettuce seed cannot be correlated directly with germination because it can be induced in conditions that retain seeds in their dormant state (Tables 2, 3). Although the role of Pfr in inducing α-galactosidase is only transient, the enzyme itself does not arise until some 3 h after irradiation, indicating that certain events set in motion by Pfr must take time to be completed.

This work is supported by NSERC of Canada grant A6352.

Received 23 September; accepted 26 November 1980.

1. Borthwick, H. A., Hendricks, S. B., Toole, E. H. & Toole, V. K. *Bot. Gaz.* **115,** 205–225 (1954).
2. Bewley, J. D., Black, M. & Negbi, M. *Nature* **215,** 648–649 (1967).
3. Bewley, J. D., Negbi, M. & Black, M. *Planta* **78,** 351–357 (1968).
4. Frédéricq, H. *Pl. Physiol.* **39,** 812–816 (1964).
5. Smith, H. *Phytochrome and Photomorphogenesis* (McGraw-Hill, London, 1975).
6. Mohr, H. *Photochem. Photobiol.* **5,** 469–483 (1966).
7. Halmer, P., Bewley, J. D. & Thorpe, T. A. *Planta* **139,** 1–8 (1978).
8. Leung, D. W. M., Reid, J. S. G. & Bewley, J. D. *Planta* **146,** 335–341 (1979).

True morphology of the Azotobacteraceae—filterable bacteria

J. Gonzalez Lopez* & G. R. Vela

Department of Biological Sciences, North Texas State University,
Denton, Texas 76203

The history of bacteriology up to the 1950s contains many references to filterable forms of bacteria but misconceptions regarding the nature of filterability eventually made the entire phenomenon an unresolvable paradox[1–4]. Although the existence of filterable azotobacter has often been reported[5–9], so much doubt in these results has been expressed[10,11] that now Bergey[12] does not even mention filterability. In 1965 we reported the resistance of soil azotobacter to γ rays and concluded that the possibility of a very small *Azotobacter* form could not be dismissed[13] and as a result of studies on survival in soil, we suggested the possible existence in soil of a still undiscovered phase of the *Azotobacter* cyst[14]. We report here the discovery of these small, soil azotobacter and provide evidence to support this claim.

The Azotobacteraceae are usually cultured in chemically defined, nitrogen-free media with an organic compound as a source of carbon. Growing cells are large, rod-shaped and pleomorphic but as cultures age, morphological variability

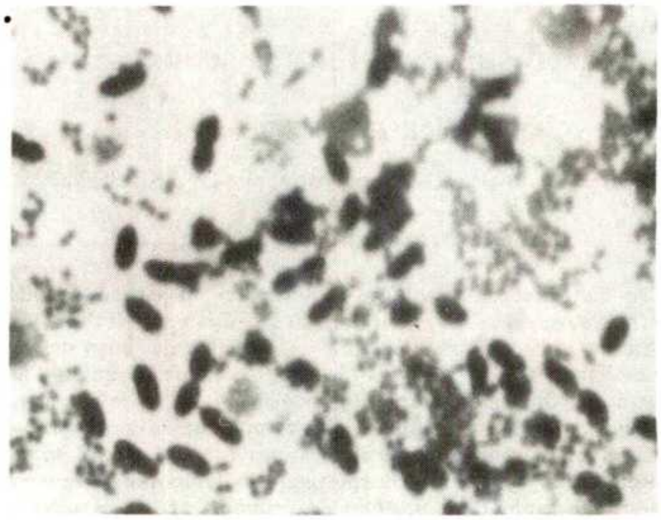


Fig. 1 Light transmission micrograph of large cells and filterable germinal cellulae of *A. vinelandii* grown in dialysed soil media. ×2,130.

---

* Permanent address: Department of Microbiology, Faculty of Pharmacy, Universidad de Granada, Granada, Spain.

0028–0836/81/070588–03$01.00

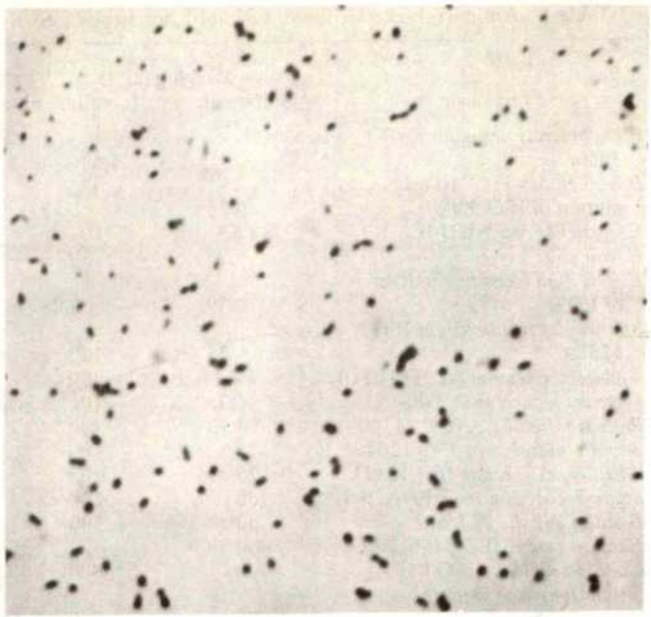


Fig. 2 Isolated germinal cellulae of *A. vinelandii* 12837 grown in dialysed soil media to a final population of $\sim 1 \times 10^7$ cellulae ml^{-1} and filtered through 0.45-μm membrane filters. ×2,220.

increases and cell debris appears, sometimes including regular, spherical or slightly elongated particles ~0.3 μm in diameter, similar to those previously reported[5,6,15]. When 10–30-day-old cultures of azotobacter were filtered through 0.45-μm membrane filters and plated on Burk's nitrogen-free medium[16], ~10% of filtrates yielded some viable cells (unpublished data). The use of different media including soil plates and different soil extract formulations did not significantly alter these findings.

In this study we pulverized dry garden soil, sifted it through gauze and placed it in dialysis tubing in 10-g amounts. Tubes were tied at both ends, placed in 50 ml of distilled water in 250-ml flasks and heated in the autoclave at 121 °C for 15 min. These conditions were established as being adequate for sterilization by the addition of sterile tryptic soy broth to selected flasks and incubation of these as sterility controls in all tests. The dialysis bags filled with soil were left in the flask cultures as a continuing source of nutrients. This growth medium remained clear, soil particles did not spill out of dialysis bags (see Fig. 1) and, when bacteria grew, they did not penetrate the bag. The dialysed soil medium was inoculated with stock cultures of *Azotobacter vinelandii* A.T.C.C. 12837 and incubated on a rotary shaker at 28 °C. After 24 h, normal cells of *A. vinelandii* were present and also many cell-like particles ~0.3 μm in diameter (Fig. 1) which passed through 0.45-μm membrane filters. Control cultures in Burk's nitrogen-free media contained only the large cells.

Filtrates were subcultured in dialysed soil medium and at each transfer, both the normal and filterable cells grew to large populations. The integrity of each filter was ascertained by adding to each azotobacter suspension a few drops of a culture of *Serratia marcescens* and plating a few drops of the unfiltered mixture and a few drops of the filtrate on tryptic soy agar. If *S. marcescens* failed to grow from the unfiltered cell mixture or if it grew from the filtrate, the test was disregarded on the grounds that it was not proved that the filter could retain bacteria. Subculture from the dialysed soil medium (unfiltered) transferred to Burk's medium yielded large cells only.

Filterable particles able to generate large cells of azotobacter have previously been called corpuscles, gondida, gonidingia, microgonidia, chromidia, autospores and endospores. We suggest that they be called germinal cellulae to describe their nature more clearly.

A. vinelandii germinal cellulae were separated from large cells by filtering 72-h-old dialysed soil cultures and centrifuging the filtrate (20,000*g*, 10 min). Microscopy revealed no large cells but many germinal cellulae (Fig. 2). When inoculated into dialysed soil media, germinal cellulae gave rise to the large cells of *A. vinelandii* (confirmed by physiological/biochemical tests) and more cellulae, but if the latter were isolated and placed in Burk's medium, they did not grow, indicating that germinal cellulae can grow only in dialysed soil medium. We suggest that the germinal cellulae are formed in large numbers ($\sim 1 \times 10^7$ per ml of culture) (Fig. 3) in the soil medium but only rarely or not at all in Burk's medium. Our experiments were repeated 32 times, with identical results. The relationship between the large cells of *A. vinelandii* 12837 and the germinal cellulae was obvious when cultures in dialysed soil media were examined at 4-h intervals for several days: germinal cellulae grew, divided and gave rise to large cells which in turn gave rise to new populations of germinal cellulae. These growth cycles were determined in two ways. First, the ratio of large cells to germinal cellulae was determined directly on stained smears of culture samples. Second, filtered and unfiltered culture samples were plated on a solid medium comprising 20 g of agar per litre of soil medium (dialysis bag discarded). The difference between the two viable cell counts represented the ratio of filterable germinal cellulae to large cells. There was good agreement in results obtained by the two methods.

Other Azotobacteraceae which gave filterable cellulae when cultured in the dialysed soil medium were: *A. vinelandii* OP, S4, S5 and S6, *Azotobacter chroococcum* AC16, *Azomonas macrocytogenes* 8702, *A. macrocytogenes* s. TM, *A. macrocytogenes* ZAM and *Beijerinckia indica*.

Our observations suggest that, in an environment similar to their natural habitat, the soil azotobacter exist as filterable cells ~0.3 μm in diameter, whereas the large cells usually seen in chemically defined laboratory cultures seem to be only a transient reproductive form.

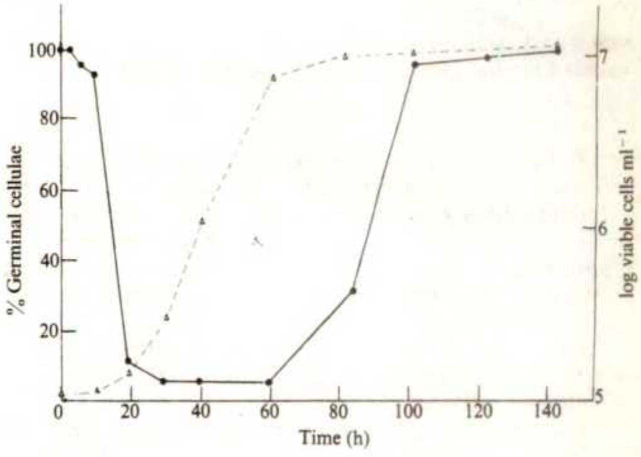


Fig. 3 *A. vinelandii* 12837 grown in dialysed soil medium at 28 °C. The continuous line shows the ratio of germinal cellulae to large cells. The broken line indicates total viable cell count on dialysed soil agar plates.

To test this hypothesis we collected 72 different soils and suspended 1-g portions of each in 10 ml of distilled water. These were agitated for 5 min, *S. marcescens* added and then filtered through 0.45-μm membrane filters. All samples inoculated into flasks of dialysed soil medium yielded azotobacter cultures whereas those inoculated into Burk's nitrogen-free medium did not. When the former cultures, containing large cells and germinal cellulae, were subcultured on Burk's nitrogen-free agar plates, they gave colonies of *A. chroococcum* and *A. vinelandii*, usually obtained from Texas soils. Replicates of this experiment always produced the same results. Three water samples from a sewage treatment plant also yielded filterable azotobacter when cultured in dialysed soil but not in nitrogen-free media. We therefore believe that azotobacter in their natural habitat exist as small, filterable cellulae incapable of growing in chemically defined, nitrogen-free media but able to grow in dialysed soil media as described here.

This work was sponsored by the Fundacion Juan March, Madrid. We gratefully acknowledge the contributions made by Henry Siu, Dr Gerald D. Cagle, Ann Collins and Robert W. Butsch who worked on the early phases of this research.

Received 17 September; accepted 25 November 1980.

1. Copper, F. B. & Petroff, S. A. *J. infect. Dis.* **43,** 200–214 (1928).
2. Frobisher, M. Jr *J. infect. Dis.* **42,** 461–472 (1928).
3. Kendall, A. I. & Walter, A. W. *J. infect. Dis.* **53,** 355–371 (1933).
4. Klieneberger-Nobel, E. *J. Hyg.* **47,** 393–395 (1949).
5. Jones, D. H. *J. Bact.* **5,** 325–341 (1920).
6. Lohnis, F. & Smith, N. R. *J. agric. Res.* **23,** 401–432 (1923).
7. Kyle, T. S. & Eisenstark, A. *Bull. Okla. agric. mech. Coll.* **48,** Biol. Sci. Ser. 5 (1951).
8. Lawrence, J. C. *Nature* **176,** 1033–1034 (1955).
9. Rubenchik, L. I. Natn. Sci. Fdn OTS 63-11076 (1963).
10. Lewis, I. M. *J. Bact.* **34,** 191–205 (1937).
11. Klieneberger-Nobel, E. *Bact. Rev.* **15,** 77–130 (1951).
12. Buchanan, R. E. & Gibbons, N. E. (eds) *Bergey's Manual of Determinative Bacteriology* 8th edn (Williams and Wilkins, Baltimore, 1974).
13. Vela, G. R. & Wyss, O. *J. Bact.* **89,** 1280–1285 (1965).
14. Aladegbami, S. L., Tsai, J. C. & Vela, G. R. *Curr. Microbiol.* **2,** 327–329 (1979).
15. Callao, V. & Hernandez Gimenez, E. *An. Edafol. Agrobiol.* **20,** 421–432 (1961).
16. Vela, G. R. & Rosenthal, R. S. *J. Bact.* **111,** 260–266 (1972).

Sulfazecin and isosulfazecin, novel β-lactam antibiotics of bacterial origin

A. Imada, K. Kitano, K. Kintaka, M. Muroi & M. Asai

Microbiological Research Laboratories, Central Research Division, Takeda Chemical Industries Ltd, Yodogawa-ku, Osaka 532, Japan

In the long history of screening for antibiotics, fungi[1,2] and actinomycetes[3–8,19] have been the only producers of β-lactam antibiotics although several phytopathogenic bacteria have been reported to produce toxins with a β-lactam structure[9,10]. We report here the first evidence that novel monocyclic β-lactam antibiotics, sulfazecin and isosulfazecin, are produced by new species of *Pseudomonas*.

Screening for β-lactam antibiotics was carried out using a selective and sensitive method involving *Pseudomonas aeruginosa* PsC^{ss} (ref. 11) and *Escherichia coli* PG8 (lacking chromosomal β-lactamase and penicillin-binding protein (PBP) 1B), which are hypersensitive to β-lactam antibiotics. Several bacterial strains producing β-lactam antibiotics were isolated from soil, mostly using agar plates acidified to *p*H 4.5. The selected bacteria were categorized into two taxonomic groups: one group represented by strain G-6302 which produced sulfazecin and another by strain SB-72310 which produced isosulfazecin. Bacterial strains of both groups are polar-flagellated Gram-negative rods; the flagellation of strain G-6302 was monotrichous and that of strain SB-72310 multitrichous. Both strains were strictly aerobic, catalase-positive, oxidase-negative, poly-β-hydroxybutyric acid-accumulating and had no nutritional requirements. Arginine dihydrolase was found in strain SB-72310, but not in strain G-6302. The G+C content of the DNA was 64 mol% in strain G–6302 and 70 mol% in strain SB-72310. On the basis of these characteristics, both strains were regarded as members of the genus *Pseudomonas*. None of the defined species of *Pseudomonas* described by Bergey[12] exhibits these particular combinations of characters. In addition, both strains were unique in their *p*H dependence for growth. They grew well at *p*H 4.0 whereas those species of *Pseudomonas* described[12] do not. However, growth was poor at alkaline *p*H: strains G-6302 and SB-72310 failed to grow at *p*H >8.3 and >8.9, respectively. We designated strain G-6302 *Pseudomonas acidophila*, and strain SB-72310 *Pseudomonas mesoacidophila*, to denote their acidophilic and moderately acidophilic properties, respectively.

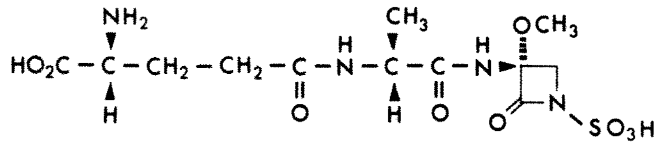


Fig. 1 Structure of sulfazecin.

Table 1 Antimicrobial activities of sulfazecin and isosulfazecin

Organism	MIC (μg ml^{-1}) Sulfazecin	Isosulfazecin
Pseudomonas aeruginosa IFO 3080	800	1,600
P. aeruginosa PsC^{ss} (sensitive mutant of IFO 3080)	0.78	0.78
Escherichia coli NIHJ JC2	12.5	100
E. coli LD2*	6.25	50
E. coli PG8 (sensitive mutant of LD2)	0.39	0.78
Salmonella typhimurium IFO 12529	6.25	100
Klebsiella pneumoniae IFO 3318	25	400
Proteus vulgaris IFO 3988	6.25	100
Proteus mirabilis ATCC 21100	3.13	100
Serratia marcescens IFO 12648	25	100
Alcaligenes faecalis IFO 13111	25	25
Staphylococcus aureus FDA 209P	200	200
Bacillus subtilis PCI 219	50	100
Bacillus cereus IFO 3466	800	400
Candida albicans IFO 0583	>800	>800
Saccharomyces cerevisiae IFO 0209	>800	>800
Aspergillus niger IFO 4066	>800	>800
Penicillium chrysogenum IFO 4626	>800	>800

Minimum inhibitory concentrations (MICs) were assayed by the twofold agar dilution method using nutrient agar containing 20 μg ml^{-1} *meso*-diaminopimelic acid. *meso*-Diaminopimelic acid was added to suffice the requirements of *E. coli* LD2 and PG8. When fungi and yeasts were the test organisms, glucose was supplemented at 1%.

* See ref. 14.

On cultivation of these bacteria in nutrient broth containing appropriate carbon and sulphur sources, antibacterial substances accumulated extracellularly. The active principles were thought to be β-lactam antibiotics because they selectively inhibited β-lactam antibiotic-hypersensitive mutants, *P. aeruginosa* PsC^{ss} and *E. coli* PG8, induced the formation of spheroplasts, and were slightly inactivated by β-lactamases.

To prepare the active principles, fermentation was carried out at 28 °C for 3 days with aeration and agitation in a 2,000-litre fermenter containing 1,200 l of medium (3% glycerol, 0.1% glucose, 0.1% sodium thiosulfate and 0.5% each of peptone, meat extract and NaCl). The antibiotics accumulated were purified using activated charcoal, anion-exchange resin and Sephadex, followed by crystallization as free acid from aqueous methanol.

The physicochemical properties of sulfazecin, produced by strain G-6302, are as follows: $C_{12}H_{20}N_4O_9S$, m.p. 168–170 °C,

0028-0836/81/070590-02$01.00

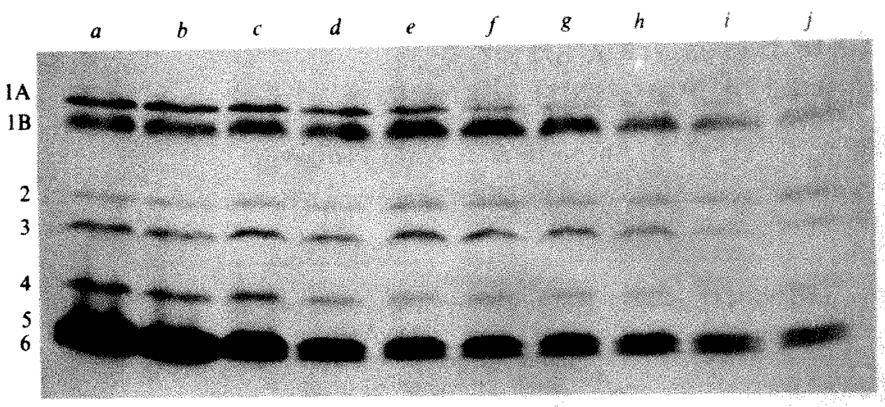
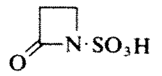


Fig. 2 Competition of sulfazecin for ^{14}C-benzylpenicillin binding to envelopes of *E. coli* LD2. The assay of PBP was performed as described[14], with slight modifications. Cell envelopes (12 mg protein ml^{-1}) of *E. coli* LD2 were preincubated with water or increasing concentrations of sulfazecin. The PBPs remaining accessible were labelled with a saturating concentration of ^{14}C-benzylpenicillin (48.2 μCi ml^{-1}, 53 mCi $mmol^{-1}$). PBPs were detected by exposure of the dried gel to an X-ray film for 26 days at −80 °C. The final concentrations of sulfazecin were (μg ml^{-1}): *a*, 0; *b*, 0.33; *c*, 0.82; *d*, 2.05; *e*, 5.12; *f*, 12.8; *g*, 32; *h*, 80; *i*, 200; and *j*, 500. Concentrations required to inhibit ^{14}C-benzylpenicillin binding by 50%, calculated from the fluorogram, were (μg ml^{-1}): PBP 1A, 11; PBP 1B, 270; PBP 2, >500; PBP 3, 150; PBP 4, 2.2; and PBP 5/6, 1.6.

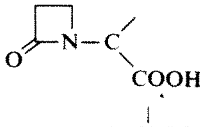
$[\alpha]_D^{25} + 82°$ ($C = 0.44$, H_2O), no UV absorption maximum above 210 nm, characteristic IR absorption bands at 1,775 cm^{-1} (β-lactam carbonyl), 1,660 cm^{-1} (amide carbonyl), 1,240 cm^{-1}, 1,040 cm^{-1} and 630 cm^{-1} (SO_3). The antibiotic gave a positive ninhydrin reaction. It was readily soluble in water but sparingly soluble in organic solvents except for polar aprotic solvents, for example, dimethylformamide. On hydrolysis with 6 M HCl at 110 °C, D-glutamic acid and D-alanine were obtained, and on mild hydrolysis with 1M HCl at 37 °C, γ-D-glutamyl-D-alanylamide and 2-oxo-3-sulphoaminopropionic acid were isolated. X-ray analysis of the intact antibiotic established the structure of sulfazecin[13]. The absolute configuration of sulfazecin was determined as 3(*R*)-3-γ-D-glutamyl-D-alanylamino-3-methoxyazetidin-2-one-1-sulphonic acid (Fig. 1) from the evidence that D-amino acids were obtained by acid hydrolysis. The name 'sulfazecin' was given to denote the fact that sulphonic acid is structurally unique in its direct link to the nitrogen atom of azetidin-2-one. Other characteristic features are a methoxyl group attached to C-3 of the azetidine ring, which is responsible for the resistance to β-lactamases (unpublished observation) and a dipeptide comprising the side chain.

The physicochemical and chromatographic properties of isosulfazecin, produced by strain SB-72310, were almost identical to those of sulfazecin, except for the optical rotation: $[\alpha]_D^{25}$ of isosulfazecin was +4.5° ($C = 1.01$, H_2O). Amino acids formed by acid hydrolysis were D-glutamic acid and L-alanine. This indicates that isosulfazecin is an epimeric isomer of sulfazecin; D-alanine in sulfazecin is replaced by L-alanine in isosulfazecin.

Sulfazecin and isosulfazecin are thus proved to be novel monocyclic β-lactam antibiotics of unique structure. The



moiety is essential to the antibacterial activity (unpublished observation) and its presence definitely differentiates sulfazecin and isosulfazecin from all known β-lactam antibiotics, including monocyclic nocardicin, that possess the



moiety. The sulphamate group (—N·SO_3H) is rarely found in nature. To our knowledge, heparin and the related *N*-sulphonated glycans are the only naturally occurring products ever found to contain this group. Although certain monocyclic β-lactams with antimetabolite activity[9,10,19] are structurally related to sulfazecin, they have neither sulphonate group at N-1 nor aminogroup at C-3 of the azetidine ring.

The antimicrobial activities of sulfazecin and isosulfazecin are shown in Table 1. Sulfazecin was active against Gram-negative bacteria and weakly active against Gram-positive bacteria. Isosulfazecin was weakly active against both Gram-negative and Gram-positive bacteria. The mutants hypersensitive to β-lactam antibiotics were extremely sensitive to both antibiotics whereas fungi and yeasts were insensitive.

Sulfazecin was also effective *in vivo*. The ED_{50} of sulfazecin administered subcutaneously was 6.98 mg per kg against mice infected intraperitoneally with *E. coli*. The toxicities of both antibiotics were extremely low; no mice died after intravenous injection of sulfazecin at 10 g per kg.

The action of sulfazecin was further examined using *E. coli* as the test organism. It was bacteriostatic at 6.25 and 12.5 μg ml^{-1}, and bactericidal at 25 μg ml^{-1} or more. The cells became filamentous when exposed to the bacteriostatic concentrations and changed to ghosts when exposed to bactericidal concentrations. Sulfazecin competed in the binding of ^{14}C-benzylpenicillin to PBPs (Fig. 2); affinity was high for PBPs 4 and 5/6 and low for 1, 2 and 3 (see Fig. 2 legend). It is interesting that the monocyclic sulfazecin, although structurally different from the bicyclic penicillins and cephalosporins, showed affinity for PBPs. The fact that the affinity for PBPs 1, 2 and 3, the apparent lethal targets in *E. coli*[15-18], was not high enough to explain the antibacterial activity suggests some alternative mode of action for sulfazecin. Further studies on the mechanism of action of sulfazecin could lead to a better understanding of the mechanism of action of β-lactam antibiotics or of biosynthesis of peptidoglycan.

We thank Dr E. Ohmura for his interest and encouragement in this work.

Received 6 October; accepted 17 November 1980.

1. Lemke, P. A. & Brannon, D. R. in *Cephalosporins and Penicillins, Chemistry and Biology* (ed. Flynn, E. H.) 370–437 (Academic, New York, 1972).
2. Kitano, K. *et al. J. Ferment. Technol.* **53,** 327–338 (1975).
3. Nagarajan, R. *et al. J. Am. chem. Soc.* **93,** 2308–2310 (1971).
4. Stapley, E. O. *et al. Antimicrob. Ag. Chemother.* **2,** 122–131 (1972).
5. Aoki, H. *et al. J. Antibiot.* **29,** 492–500 (1976).
6. Brown A. G. *et al. J. Antibiot.* **29,** 668–669 (1976).
7. Kahan, J. S. *et al. J. Antibiot.* **32,** 1-12 (1979).
8. Brown, A. G., Corbett, D. F., Eglington, A. J. & Howarth, T. T. *JCS chem. Commun.* 523–525 (1977).
9. Stewart, W. W. *Nature* **229,** 174–178 (1971).
10. Taylor, P. A., Schnoes, H. K. & Durbin, R. D. *Biochim. biophys. Acta* **286,** 107–117 (1972).
11. Kitano, K., Kintaka, K. & Nakao, Y. *J. Ferment. Technol.* **54,** 696–704 (1976).
12. Doudoroff, M. & Palleroni, N. J. in *Bergey's Manual of Determinative Bacteriology* 8th edn (eds Buchanan, R. E. & Gibbons, N. E.) 217–243 (Williams & Wilkins, Baltimore, 1974).
13. Kamiya, K. *et al.* (in preparation).
14. Nozaki, Y., Imada, A. & Yoneda, M. *Antimicrob. Ag. Chemother.* **15,** 20–27 (1979).
15. Tamaki, S., Nakajima, S. & Matsuhashi, M. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5472–5476 (1977).
16. Spratt, B. G. *Nature* **274,** 713–715 (1978).
17. Suzuki, H., Nishimura, Y. & Hirota, Y. *Proc. natn. Acad. Sci. U.S.A.* **75,** 664–668 (1978).
18. Iwaya, M. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2980–2984 (1977).
19. Scannel, J. P. *et al. J. Antibiot.* **28,** 1–6 (1975).

Mite faeces are a major source of house dust allergens

E. R. Tovey, M. D. Chapman & T. A. E. Platts-Mills*

Division of Immunology, Clinical Research Centre, Watford Road,
Harrow, Middlesex HA1 3UJ, UK

The association between house dust allergy and asthma has long been recognized, and it has been demonstrated that a major allergen in house dust is related to the presence of mites of the genus *Dermatophagoides*[1]. Using extracts of mite culture for skin testing, as many as 10% of the population and up to 90% of allergic asthmatics give positive immediate reactions[2]. Although mites may occasionally become airborne during bed-making[3], it has also been demonstrated that they 'secrete or excrete' some allergen[1]. Recently, we have shown that up to three-quarters of the serum IgE antibodies to mites are directed against a major allergen—antigen P_1 (molecular weight 24,000)[4]. Using a radioimmunoassay it is possible to measure the concentration of this glycoprotein in both dust samples and mite cultures. These measurements, which are reported here, show that more than 95% of the allergen accumulating in mite cultures is associated with faecal particles.

When the components of mite culture were separated, whole mites, cuticles and mite faeces contained antigen P_1 but eggs contained very little (Table 1). The number of faecal particles in cultures seemed to be very high from direct microscopy. Using a mean value of 20 faeces per day, we estimated that after 1 month faeces would account for >95% of the antigen P_1 in cultures. These particles are roughly spherical and have a relatively smooth surface (Fig. 1). They are 10–40 μm in diameter and their size is directly proportional to the size of the mites producing them, which vary in length from 170 to 350 μm ($r = 0.91$, $P < 0.001$). The antigen P_1 content of faecal particles, mite culture or house dust elutes very rapidly in saline, whereas the allergen content of live mites is only very slowly eluted (Fig. 2).

Table 1 Quantity of antigen P_1 in mite culture components

	Culture components	ng antigen P_1 per 100 components	% Of total antigen P_1
Mites	Female	185	0.6
	Male	105	
	Immature	75	
Cuticles		29	0.4
Eggs		0.4	–
Faeces	Mixed size (22 ± 6 μm)	12	99
	Large (31 ± 8 μm)	17	
	Small (17 ± 4 μm)	3	

Antigen P_1 content of separated components of mite culture. Live mites (Bencard) were grown on Oxoid liver powder for 1 week at 24 °C and 75% relative humidity. Individual components were separated, homogenized, extracted with borate-buffered saline and assayed for P_1 content by double antibody inhibition radioimmunoassay[4]. The contribution of the components to the total P_1 content of culture is based on a mean lifetime for a mite of 3 months involving three changes of cuticle and an estimated rate of defaecation of 20 faeces per day. Two methods were used to measure the rate of defaecation. When the mean time for the passage of coloured food through the mite gut (5 h) was multiplied by the mean number of faeces produced by an isolated fully fed mite (mean 8 faeces, range 6–11), a rate of 40 faeces per mite per day was obtained. Alternatively, counting the number of faeces produced by mites fed on large washed lumps of culture media after 6 days in culture gave a mean rate of defaecation of only 6 faeces per day in these suboptimal conditions. Faecal sizes are means ± s.d.

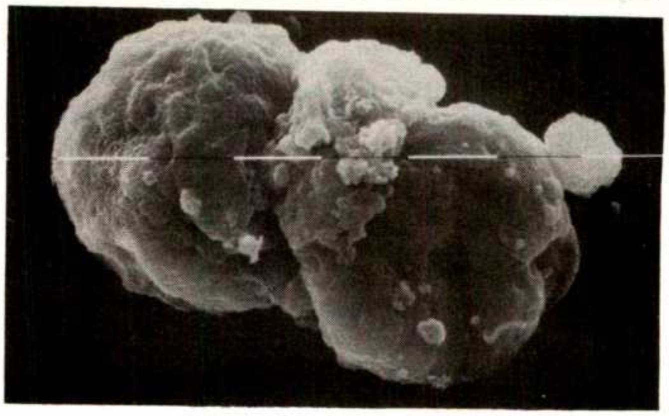


Fig. 1 Electron scanning micrograph of two mite faecal balls. Mite faeces range from 10 to 40 μm in diameter with a mean of 22 ± 6 μm s.d. For a similar species it has been reported[9] that the faeces are produced by compacting three to five foodballs covered in a peritrophic membrane. Scale bar, 10 μm.

As most particles deposited on the nasal mucosa are swallowed within 10 min, rapid elution of proteins is probably essential for allergenic activity[5]. It is not uncommon for house dust samples to contain over 1,000 mites and 30 μg of antigen P_1 per gramme. Based on the values in Table 1, such samples probably contain over 250,000 faeces per gramme of dust. Our further studies have shown that antigen P_1 is only present in the air of houses during domestic activity and this allergen is associated with particles >10 μm in diameter[6]. In general, particles of this size would be deposited on the mucosa of the nose and oropharynx and only occasional particles would be expected to enter the lungs[7]. In keeping with this, most dust-allergic asthmatic patients are not aware of a direct relationship between dust exposure and the onset of bronchospasm. The concentration of allergen in mite faeces is very high, approximately 10 mg antigen P_1 per ml, assuming 0.1 ng of allergen in a sphere of 20 μm diameter. These particles, whose physical properties are similar to those of pollen grains, would seem to be a very effective way of carrying proteins to the nasal mucosa. In addition, the few particles entering the lungs would be expected to cause localized inflammatory responses because of the high concentrations of allergen[8]. Our results suggest that the relationship between dust mite allergy and bronchial asthma is best

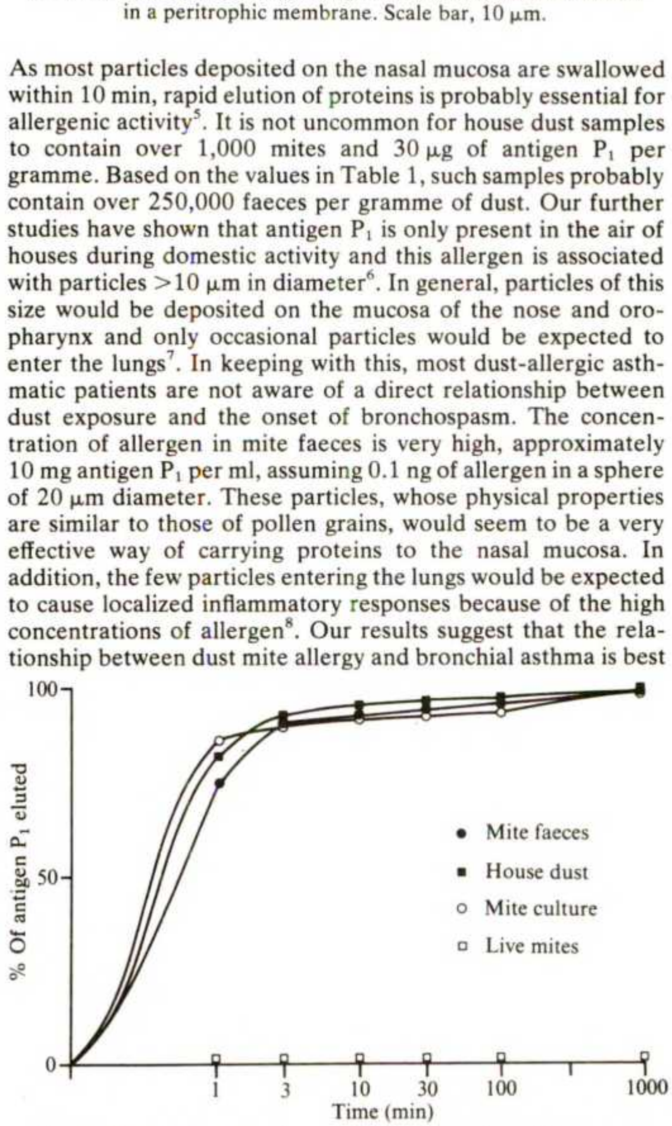


Fig. 2 The elution of antigen P_1 from faeces, live mites, house dust and culture material. Mite faeces or live mites mounted on Cellotape (1 × 5 mm) were contained in a section of capillary tubing and eluted with saline using a peristaltic pump (25 μl min^{-1}). At the end of the elution period 11 of the 60 mites were still mobile and none was observed to defaecate during the elution. Samples of house dust and culture material were repeatedly eluted in filter funnels under vacuum (30 ml, Scinterglass 4, 5 μm pore size). With house dust, culture material and faeces, the quantities of P_1 are expressed as a % of the total eluting over 16 h. With live mites the quantity of P_1 is expressed as a percentage of the quantity obtained by homogenizing the mites.

---

* To whom correspondence should be addressed.

0028-0836/81/070592-02$01.00

explained if inflammatory responses caused by faecal particles have a cumulative effect on the bronchi.

We thank the Asthma Research Council who provided financial support for E. R. T. and M. D. C., and John Clarke for help with electron microscopy.

Received 18 August; accepted 15 December 1980

1 Voorhorst, R , Spieksma Boezeman, M I A & Spieksma, F Th M *Allergie Asthma* **10**, 329–334 (1964).
2 Morrison-Smith, J , Disney, M E , Williams, J P. & Goels, Z A *Br med J* **i**, 723–726 (1969)
3 Cunnington, A M & Gregory, P H *Nature* **217**, 1271–1272 (1968)
4 Chapman, M D & Platts-Mills, T. A. E *J Immun* **125**, 587–592 (1980)
5 Marsh, D G. in *The Antigens* Vol 3 (ed. Sela, M) 271–350 (Academic, New York, 1975)
6 Tovey, E R , Chapman, M D & Platts-Mills, T A. E (in preparation)
7 Harper, G J & Morton, J D *J Hyg , Camb* **51**, 374–385 (1953)
8 Anston, F. K & Orange, R P *Am Rev resp Dis* **112**, 423–436 (1975).
9 Brody, A R., McGarth, J C & Wharton, G W *N Y. ent Soc* **80**, 152–177 (1972)

Active antitoxic immunization by a diphtheria toxin synthetic oligopeptide

F. Audibert[*], M. Jolivet[*], L. Chedid[*], J. E. Alouf[†], P. Boquet[†], P. Rivaille[‡] & O. Siffert[§]

[*] Groupe de Recherche No. 31 du CNRS, Immunothérapie Expérimentale, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cédex 15, France
[†] Unité des Antigènes Bactériens (ERA-CNRS 74), Institut Pasteur, Paris, France
[‡] LA-CNRS 163, CHU St-Antoine, 27 rue de Chaligny, 75012 Paris, France
[§] Unité de Chimie Organique, Institut Pasteur, Paris, France

Diphtheria toxin (DT) is a single polypeptide chain of molecular weight 62,000 with two disulphide bridges[1]. Immunization against diphtheria rests on the stimulation of antibodies against detoxified toxin which also combine with the native toxin. Because the antibodies differ functionally from each other, however, only some of them are able to neutralize toxicity[2–4]. We have therefore set out to synthesize part of the amino acid sequence of the toxin whose function as a stimulator of antibodies might be less ambiguous, and have chosen the loop of 14 amino acids subtended by the disulphide bridge nearer the NH_2 terminus of the molecule (Fig. 1). There is reason to think that this loop may be involved in the toxicity and immunological specificity of the molecule[1,5]. We report here our finding that the tetradecapeptide (residues 188–201), when linked covalently with two different carriers, will elicit in guinea pigs antibodies which not only bind specifically with the toxin but neutralize its dermonecrotic and lethal effects. To our knowledge these results constitute the first example of successful active immunization against a lethal bacterial toxin using a synthetic antigen.

186 190 195 201
-Cys-Ala-Gly-Asn-Arg-Val-Arg-Arg-Ser-Val-Gly-Ser-Ser-Leu-Lys-Cys-

Fig. 1 Sequence of the diphtheria toxin loop within the NH_2-terminal disulphide bridge as inferred from the complete sequence of fragment A (ref. 14) and the data reported for the NH_2-terminal stretch of fragment B (ref. 15).

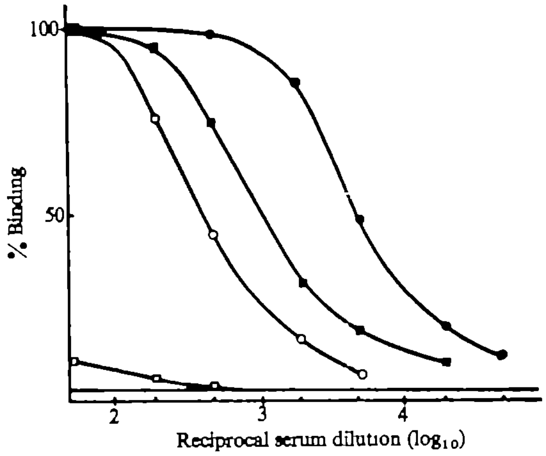


Fig. 2 Binding of radiolabelled diphtheria toxin by anti-SDP and anti-STDP sera. DT was iodinated by the chloramine-T technique. Using appropriate dilutions, 98% of the radioactivity present in the antigenic preparation after labelling was bound by a hyperimmune anti-DT horse serum (gift from M. Mazert). Serial dilutions of anti-toxoid and anti-peptide sera were tested for their capacity to bind ^{125}I-labelled diphtheria toxin. Percentages were obtained in each tube from the formula: c.p.m. found in precipitate ×100/total c.p.m. ●, Anti-toxoid serum corresponding to the pooled sera of group 1 (see Table 1 legend); ■, anti-toxoid serum (pooled sera of group 2); ○, anti-SDP (pooled sera of group 5); □, anti-STDP serum (highest responder of group 6); —, normal rabbit serum.

The sequences 194–201, 192–201 (synthetic decapeptide SDP), 190–201, 188–201 (synthetic tetradecapeptide STDP) and 186–201 were prepared by solid-phase methodology using *m*-nitrobenzhydrylamine resin and benzhydrylamine resin. The first polymer retains the peptide after hydrogen fluoride (HF) treatment to remove protecting groups of the side chains of the amino acids[6]. Despite certain gaps, such as those caused by destruction of cysteine, the elongation of the sequences can be followed by amino acid analysis performed on the peptidyl resin[7]. With the second polymer, HF treatment cleaves peptides in their carboxylamide terminal form[8]. Peptides were purified by gel filtration and proved homogeneous by TLC. Their amino acid composition agreed with that obtained theoretically. Details of the different syntheses will be published elsewhere.

Octa-, deca- (SDP), dodeca-, tetradeca- (STDP) and hexadecapeptides linked to the *m*-nitrobenzyhydrylamine resin were tested for their ability to combine with anti-DT antibodies. These sera were raised in horses and rabbits, and the immunoglobulins purified and labelled with ^{125}I. Whereas SDP seemed to be the smallest structure capable of binding to antibodies, this property increased strikingly with the hexadecapeptide. Because free SDP and STDP were the only compounds available in sufficient amounts, they were covalently linked to bovine serum albumin (BSA) by their amino groups using glutaraldehyde and administered to guinea pigs in conditions described in Table 1 legend. The resulting sera were tested by passive haemagglutination. Using the SDP–BSA conjugate, only one out of five animals gave detectable response, whereas all but one of the animals challenged with STDP–BSA gave significant

Table 1 Individual response of guinea pigs to toxoid, SDP and STDP

Immunization	Passive haemagglutination titre Mean	Individual values
Toxoid:		
(1) 3 Lf+FCA	13.8	14 - 14 - 14 - 14 - 13
(2) 3 Lf	11.8	12 - 12 - 12 - 12 - 11
(3) 0.3 Lf	3.8	7 - 4 - 4 - 2 - 2
(4) 0.03 Lf	0.2	1 - 0 - 0 - 0 - 0
(5) SDP–BSA	1.6	8 - 0 - 0 - 0 - 0
(6) STDP–BSA	4.67	10 - 5 - 5 - 4 - 4 - 0

[1] Groups of guinea pigs were injected in each hind footpad with 0.1 ml of the immunizing preparations and 30 days later re-injected in the same conditions. Data represent the $\log_2$ of passive haemagglutination titres of individual sera collected 3 weeks after the second injection, using sheep red blood cells coated with diphtheria toxin by glutaraldehyde in experimental conditions to be described elsewhere. Groups are the following: guinea pigs immunized by diphtheria formalin-toxoid: (1) 3 Lf in FCA; (2) 3 Lf in saline; (3) 0.3 Lf in saline; (4) 0.03 Lf in saline; guinea pigs immunized by peptide conjugates: (5) SDP–BSA 1 mg in FCA; (6) STDP–BSA 1 mg in FCA. 1 Lf = 2.5 μg protein; 1 MLD = 66 ng protein; 1 MRD = 0.1 ng protein.

0028-0836/81/070593-02$01 00

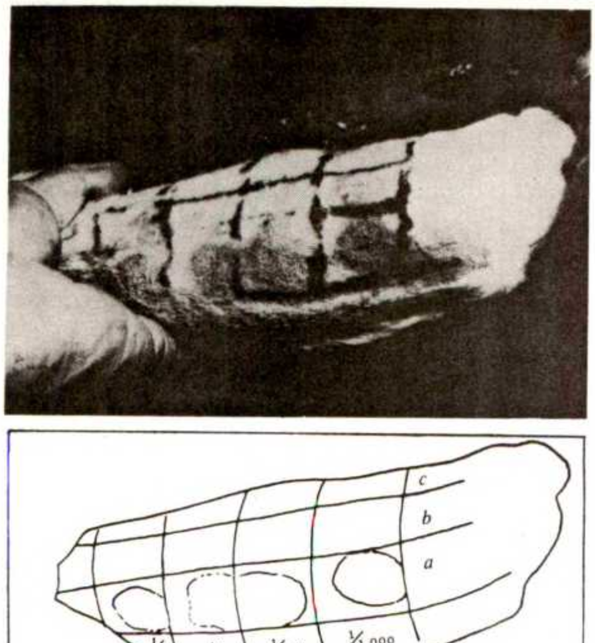


Fig. 3 Sera were diluted 2-, 10-, 100- and 1,000-fold in phosphate-buffered saline. They were mixed (1 : 1 v/v) with a 1,000 MRD ml^{-1} dilution of DT in calf serum. After incubation for 20 min at room temperature, 0.2 ml was injected intradermally in guinea pigs. *a*, Normal guinea pig serum; *b*, anti-STDP serum highest responder of group 6; *c*, anti-toxoid serum (pooled sera of group 2). Whereas positive dermonecrotic responses were observed with normal serum (*a*), responses were uniformly negative at all dilutions of *b* and *c* sera tested.

levels of antibodies. The titre ($\log_2$) of these pooled sera was 7.67 whereas pooled sera of guinea pigs immunized with 0.3 Lf of toxoid gave a titre of 5.15. (However, the latter preparation, unlike the former, was not injected with Freund's adjuvant.) No haemagglutination was observed with undiluted STDP antisera incubated with uncoated erythrocytes as controls. Radioimmunoassay gave results in agreement with those of the passive haemagglutination test (Fig. 2). Highly purified diphtheria toxin used in the present study was 90% in the 'nicked' form as evidenced by SDS-polyacrylamide gel electrophoresis. Previous incubation with DT inhibited the haemagglutinating and binding activities of anti-peptide sera.

More interesting is the finding that anti-STDP antibodies neutralized the dermonecrotic activity of DT. Necrosis caused by 100 minimum reactive doses (MRD), and both necrosis and erythema caused by 10 MRD, were totally inhibited by 10-fold diluted pooled sera (not shown) or even by a 1,000-fold dilution of the serum with the highest binding and agglutinating titres (Fig. 3*b*). Preliminary experiments on protection against lethal activity of the toxin showed that 0.5 ml of the latter anti-STDP serum protected guinea pigs against two minimal lethal doses (MLD).

Further experiments, which will be detailed elsewhere, also indicate that DT binding antibodies could be obtained by STDP coupled to the synthetic branched copolymer multi-poly-DL-alanyl-poly-L-lysine (abbreviated A--L). Moreover, DT binding antibodies have also been obtained by administering a synthetic adjuvant MurNAc-L-Ala-D-isoGln(MDP)[9,10] with either STDP–BSA or STDP–A--L in saline instead of Freund's complete adjuvant (FCA). Previous investigators used synthetic antigens containing a portion of sequences of various proteins to study the structural parameters of immunogenicity[11,12]. A pioneering example is that of lysozyme, for which the loop region between Cys 64 and Cys 80 was prepared synthetically and when coupled to poly-DL-alanyl-poly-L-lys, successfully elicited antibodies reactive with native lysozyme[13]. The work developed here could provide useful tools for a closer investigation of the molecular interaction between DT and its fragments with sensitive cells. Antibodies directed against a limited portion of the toxin molecule such as the NH_2-terminal loop should present a reduced heterogeneity. In addition, as the antigen was obtained by chemical synthesis, the biological activity of analogues can also be investigated.

These experiments and those in progress suggest the possibility of developing a completely synthetic antidiphtheria vaccine. Such a model could be extended to other infective microorganisms whenever their immunogenic structures can be defined and synthesized.

We thank A. M. Pappenheimer Jr and W. B. Vernon for the useful discussions which initiated the present work, Dr Raulais for the amino acid analysis and Ms Bassot for technical assistance. This work has been supported by INSERM grant CRL. 80.10.02.

Received 15 August; accepted 1 December 1980.

1. Pappenheimer, A. M. Jr *A. Rev. Biochem.* **46,** 69–94 (1977).
2. Raynaud, M. in *Mechanisms of Hypersensitivity* (eds Shaffer, J. H., Lo Grippo, G. A. & Chase, M. W.) 27–46 (Little Brown, Boston, 1959).
3. Raynaud, M. in *Antibodies to Biological Active Molecules* Vol. 1 (ed. Cinader, B.) 197–251 (Pergamon, London, 1966).
4. Pappenheimer, A. M. Jr, Uchida, T. & Harper, A. A. *Immunochemistry* **9,** 891–906 (1972).
5. Collier, R. J. *Bact. Rev.* **39,** 54–85 (1975).
6. Drouault, R., Valero, D. & Rivaille, P. *Proc. XVth Eur. Peptide Symp.*, 607–609 (Wroclaw University Press, 1979).
7. Scotchler, J., Lozier, R. & Robinson, A. B. *J. org. Chem.* **35,** 3151–3152 (1970).
8. Pietta, P. G., Cavallo, P. F., Marshall, G. & Pace, M. *Gaz. chim. Ital.* **103,** 483–489 (1973).
9. Chedid, L., Audibert, F. & Johnson, A. G. *Prog. Allergy* **25,** 63–105 (1978).
10. Parant, M. *Springer Semin. Immunopath.* **2,** 101–118 (1979).
11. Alouf, J. E. in *Immunologie* (ed. Bordet, P.) 209–270 (Flammarion Médecine-Sciences, Paris, 1978).
12. Arnon, R. *Colloques INSERM, Paris* **72,** 11–24 (1977).
13. Arnon, R., Maron, E., Sela, M. & Anfinsen, C. B. *Proc. natn. Acad. Sci. U.S.A.* **68,** 1450–1455 (1971).
14. De Lange, R. J., Williams, L. C., Drazin, R. E. & Collier, R. J. *J. biol. Chem.* **254,** 5838–5842 (1979).
15. Falmagne, P., Lambotte, P., Capiau, C., Ruysschaert, J. M. & Dirkx, J. *Toxicon* **17,** Suppl. 1, 46 (1979).

Thyroid growth-blocking antibodies in primary myxoedema

H. A. Drexhage*†, G. F. Bottazzo*, L. Bitensky†, J. Chayen† & D. Doniach*

* Department of Immunology, Middlesex Hospital Medical School, London W1P 9PG, UK
† Division of Cellular Biology, Kennedy Institute of Rheumatology, London W6 7DW, UK

Adult primary myxoedema is usually due to an autoimmune thyroiditis characterized by progressive shrinking of the thyroid gland, loss of epithelium, dense infiltration by sensitized lymphocytes and plasma cells with final replacement of the gland by a fibrous scar[1]. Antibodies directed against thyroglobulin (TgHA) and microsomal antibodies (McHA), detectable years before the onset of hormonal failure, and cell-mediated immune mechanisms contribute to the pathogenesis[2]. When 90% of the gland is destroyed, the secretion of thyroid hormones (T_3/T_4) falls below normal needs and pituitary thyrotrophs react by producing over 100 times the normal amount of thyroid-stimulating hormone (TSH)[3] which should normally lead to regrowth of the gland as happens in goitrous Hashimoto thyroiditis[4]. The inability of the thyroid in primary myxoedema to respond to the trophic action of TSH suggested that blocking antibodies exist in this disease which compete with TSH for its receptors[5,6]. The trophic effect of TSH can be assayed *in vitro* by measuring DNA-synthesis by Feulgen cytophotometry[7]; elevated pentose-shunt oxidative activity correlates with DNA synthesis[8]. Here we report that immunoglobulins from patients with adult, primary myxoedema block the trophic effect of TSH, as assessed by this *in vitro* system. This is in marked contrast to the growth stimulation induced by immunoglobulins from patients with thyrotoxic Graves' disease with prominent goitres[7].

0028-0836/81/070594-03$01.00

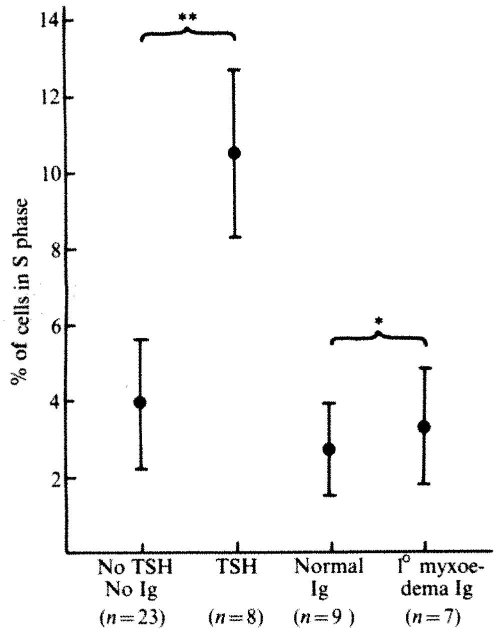


Fig. 1 DNA synthesis in thyroid explants measured by Feulgen densitometry. Guinea pig thyroid segments were exposed for 5 h to TSH, to immunoglobulin (Ig) of normal subjects, to immunoglobulin from primary myxoedema patients, or were cultured without any of these. **, $P<0.01$; *, not significant.

Immunoglobulins were prepared by ammonium sulphate precipitation from the sera of 10 middle-aged female patients with primary myxoedema. All had impalpable glands and, despite exaggerated TSH responses to exogenous thyrotropin-releasing hormone, they produced negligible secretion of thyroid hormones. Nine had detectable thyroid antibodies (McHA titre 10^2–160^2; TgHA titre 20–>5,000). Most of these patients were on replacement therapy with thyroxine at the time of testing. As controls, immunoglobulins from subjects with the following conditions were studied: (1) Goitrous Hashimoto thyroiditis (three female patients), with even higher titres of thyroid antibodies (McHA 80^2–$1{,}250^2$; TgHA 20–>5,000). (2) Graves' thyrotoxicosis, two patients with goitres of estimated weight >60 g and two without thyroid enlargement (<30 g). Thyroid growth-stimulating antibodies had been demonstrated in the former, but not the latter pair[7]. (3) A case of congenital hypothyroidism (athyreotic cretinism) due to failure of the descent of the thyroid gland, in which the myxoedema, in the presence of high levels of TSH, is due to lack of functioning thyroid tissue. In addition, normal immunoglobulin from a pool of 20 healthy subjects and from two members of the laboratory staff were tested.

For each comparative study, five segments of a single guinea pig thyroid gland were maintained at 37 °C for 5 h in Trowell's non-proliferative organ culture[9] with TSH (0.3 μU ml^{-1}; MRC Research Standard A) in the medium. The sixth segment was maintained without TSH. For four of the segments exposed to TSH, the medium also contained a standard concentration of the relevant immunoglobulin preparation (125 μg ml^{-1}) for the entire 5-h period. This immunoglobulin concentration was shown to be optimal for stimulating DNA synthesis[7]. At the end of the 5-h period, the segments were chilled, sectioned and reacted either by the Feulgen reaction[7] or for glucose 6-phosphate dehydrogenase (G6PD) activity with neotetrazolium chloride as the final hydrogen acceptor and phenazine methosulphate as intermediary carrier[10]. The reaction products were measured in 50–100 individual thyroid-follicle cells by scanning and integrating microdensitometry[11] (×100 oil immersion objective; scanning spot 0.2-μm diameter in the plane of the section; λ 550 nm for the former and 585 nm for the latter). By Feulgen cytophotometry, DNA synthesis (S phase and G_2) was defined by the presence of at least 2.8 times the C value, so excluding the normal distribution of cells in G_0 or G_1 close to the $2C$ value. ($2C$ is the normal diploid DNA content of resting cells, which doubles to $4C$ before mitosis.)

TSH stimulated DNA synthesis, with 10.3% (±2.5% s.d.) of the cells in S phase compared with 3.9% (±1.7%) in segments maintained in the absence of TSH (Fig. 1). An equivalent stimulation of G6PD activity (145% ± 15% of control value) by TSH was also recorded. The immunoglobulin derived from each of the 10 patients with primary myxoedema blocked the stimulation of DNA synthesis, depressing the number of cells found in S phase down to 28–88% of the value found with TSH alone (Fig. 2). The immunoglobulin from seven of these patients also diminished the TSH-induced stimulation of G6PD activity. In contrast, the immunoglobulin from most of the other cases enhanced the stimulatory effects of TSH in both tests, or at least did not counteract the stimulation induced by TSH. The only exception was obtained with the immunoglobulin from a patient with Graves' thyrotoxicosis but with no palpable goitre, which showed slight blocking of DNA synthesis but no blockade of the G6PD stimulation.

In general, the results of the nucleic acid cytophotometry correlated well with those of the G6PD assay ($r=0.59$; $P<0.01$). However, with a few of the immunoglobulins tested, discrepancies were found between the two tests, probably due to the involvement of the pentose shunt in metabolic functions other than growth[12].

When they were tested alone in this system, none of the immunoglobulins that blocked the effect of TSH affected DNA synthesis (Fig. 1). There was no correlation between the growth-blocking potentiality of any given immunoglobulin preparation and the titre of microsomal antibody ($r=0.22$; $P>0.05$).

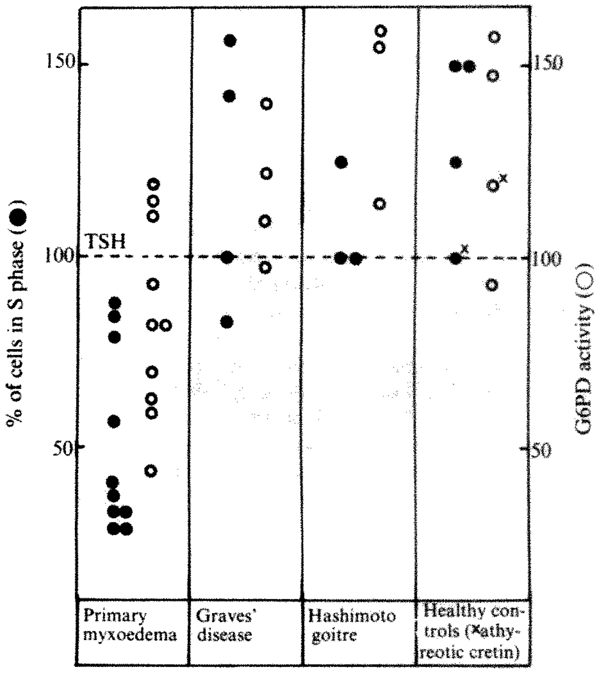


Fig. 2 Effect of the addition of immunoglobulins from several groups of patients on TSH-stimulated DNA synthesis and pentose-shunt activity in guinea pig thyroid tissue. The DNA synthesis (% of cells in S phase, ●) and the pentose-shunt activity (G6PD activity per 100 cells, ○) are expressed as a percentage of the results found with TSH alone within each assay (dashed line = TSH level = 100 per cent).

In support of our results, it has been suggested that spontaneous hypothyroidism preceding or following an attack of thyrotoxicosis could be due to blocking antibodies[5,13]. Moreover, previous studies have shown that occasionally immunoglobulins from patients with Graves' disease[14] and from two patients with myxoedema[6,15] had blocking potential in that they inhibited the stimulation of adenylate cyclase induced by TSH or by other Graves' immunoglobulins. Adenylate cyclase-inhibiting antibodies have recently been reported to cross the placental barrier, giving rise to transient hypothyroidism in the newborn[15]. A similar phenomenon explains the report[16] in which a mother treated for primary myxoedema gave birth to several babies with temporary hypothyroidism. Antibodies able to block TSH-induced stimulation of growth rather than TSH-induced hormone synthesis were less likely to be responsible for the transient hypothyroidism in this family because the thyroid gland was of normal size in one baby who died[16]. However, antibodies influencing thyroid growth are likely to cross the placenta as transient neonatal thyrotoxicosis is often accompanied by goitre[17]. Blocking antibodies have been demonstrated in other autoimmune conditions—acetylcholine receptor antibodies in myasthenia gravis[18], insulin receptor antibodies in insulin-resistant diabetes with acanthosis nigricans[19] and parathyroid hormone receptor antibodies in some cases of renal failure[20]. Our study indicates the existence of antibodies which block the trophic effects of TSH. These antibodies are probably distinct from those that block TSH-induced hormone synthesis through the adenylate cyclase pathway. Thyroid atrophy, together with loss of function, might be caused by the simultaneous presence of these two types of blocking antibodies in primary myxoedema.

We thank Ms L. Hammond for technical assistance and Miss P. C. Lamb for typing the manuscript. H.A.D. was supported by a Royal Society Exchange Fellowship. Work was also supported by a MRC grant.

Received 17 July; accepted 16 December 1980.

1. Bonnyns, M. & Bastenie, P. A. *J. clin. Endocr. Metab.* **27,** 849–854 (1967).
2. Doniach, D. & Roitt, I. M. in *Clinical Aspects of Immunology* (eds Gell, P. G. H., Coombs, R. R. A. & Lachmann, P. J.) 1355 (Blackwell, Oxford, 1975).
3. Ridgeway, E. C., Weintraub, B. D. & Maloof, F. *J. clin. Invest.* **53,** 895–903 (1974).
4. Doniach, D., Bottazzo, G. F. & Russel, R. C. G. *Clinics Endocr. Metab.* **8,** 63–80 (1979).
5. Lamberg, B. A. *Clinics endocr. Metab.* **8,** 3–20 (1979).
6. Endo, K. *et al. J. clin. Endocr. Metab.* **46,** 734–739 (1978).
7. Drexhage, H. A., Bottazzo, G. F., Bitensky, L., Chayen, J. & Doniach, D. *Lancet* **ii,** 287–292 (1980).
8. Coulton, L. A. *Histochemistry* **50,** 207–215 (1977).
9. Trowell, O. A. *Expl Cell Res.* **16,** 118–147 (1959).
10. Chayen, J., Bitensky, L. & Butcher, R. G. *Practical Histochemistry* (Wiley, London, 1973).
11. Chayen, J. in *Biochemical Mechanisms of Liver Injury* (ed. Slater, T. F.) 257-291 (Academic, New York, 1978).
12. Chayen, J., Bitensky, L., Butcher, R. G. & Altman, F. P. *Adv. Steroid Biochem. Pharmac.* **4,** 1–56 (1976).
13. Christy, J. H. & Morse, R. S., *Am. J. Med.* **62,** 291–296 (1977).
14. Orgiazzi, J., Williams, D. E., Chopra, I. J. & Solomon, D. H. *J. clin. Endocr. Metab.* **42,** 341–354 (1976).
15. Matsuura, N. *et al. New Engl. J. Med.* **303,** 738–741 (1980).
16. Goldsmith, R. E., McAdams, A. G., Larsen, P. R., McKenzie, M. & Hess, E. V. *J. clin. Endocr. Metab.* **37,** 265–275 (1973).
17. Munro, D. A., Dirmikis, S. M., Humphries, H., Smith, T. & Broadhead, G. D. *Br. J. Obstet. Gynaec.* **85,** 837–843 (1978).
18. Lennon, V. A. in *Clinical Immunology Update* (ed. Franklin, E. C.) 259–289 (Elsevier, New York, 1979).
19. Flier, J. S., Kahn, C. R. & Roth, J. C. *New Engl. J. Med.* **300,** 413–419 (1979).
20. Juppner, H., Bialasiewicz, A. A. & Hesch, R. D. *Lancet* **ii,** 1222–1224 (1978).

Block of acetylcholine-activated ion channels by an uncharged local anaesthetic

D. C. Ogden, S. A. Siegelbaum & D. Colquhoun

Department of Pharmacology, University College London,
Gower Street, London WC1E 6BT, UK

It is now thought that amine local anaesthetic compounds (procaine, lignocaine and related molecules) depress electrical activity in nerve and muscle cells by binding to sites within ion channels and blocking current flow. Such mechanisms have been proposed to account for the effects of these local anaesthetics on both the voltage-dependent sodium current[1-3] and the postsynaptic acetylcholine (ACh)-activated ionic current[4-7]. Recently, strong evidence for block of ion channels by cationic drug molecules has been obtained by recording current from single ACh-activated channels in the presence of permanently charged quaternary derivatives of lignocaine[8]. Most amine local anaesthetic compounds are, however, weak bases, present in both charged and uncharged forms at physiological *p*H, and some question remains as to whether a charged group is essential for blockade of ion channels[1,9,10]. To resolve this question, we studied the action of the uncharged local anaesthetic benzocaine (ethyl-4-aminobenzoate) on postsynaptic ACh- activated endplate current and extrajunctional single channel current of frog muscle. We report here evidence that strongly suggests that benzocaine blocks ACh-activated ion channels.

The effect of benzocaine on spontaneous miniature endplate currents (m.e.p.cs) is illustrated in Fig. 1. In the absence of drug (Fig. 1*a*), the m.e.p.c. decays along a single exponential time course with time constant τ_c. After addition of 300 μM benzocaine (Fig. 1*b*), the peak m.e.p.c. amplitude is reduced and the current decline becomes biphasic, with an initial rapid component followed by a slow tail. This decay can be fitted well by the sum of two exponential components, with time constants faster, τ_f, and slower, τ_s, than the control time constant, τ_c. On increasing the drug concentration to 500 μM (Fig. 1*c*), there is a further reduction in peak m.e.p.c. amplitude, the fast exponential current component becomes more rapid, and the slow decay phase becomes slower.

Analysis of ACh-induced current fluctuations in the absence and presence of benzocaine reveals similar effects on endplate current kinetics to those seen with m.e.p.cs (Fig. 2). Figure 2*a* shows a power spectrum of ACh-induced current noise in the absence of benzocaine. The spectrum is well described by a single lorentzian function, characterized by a cut-off frequency, f_c, which is related to the time constant, τ, by $\tau = (2\pi f_c)^{-1}$. In 200 μM benzocaine, Fig. 2*b*, the power spectrum is composed of two lorentzian components with cut-off frequencies corresponding to fast and slow time constants, τ_f and τ_s. These time constants from noise analysis are in good agreement with the time constants obtained from fitting m.e.p.c. decay (within experimental error, when normalized to the same temperature).

The results with benzocaine resemble the reported action of charged local anaesthetic compounds[4-6,11] and suggest, as an initial hypothesis, the following mechanism for open channel block[12]:

$$R \underset{\alpha}{\overset{\beta'}{\rightleftharpoons}} R^* \underset{k_b}{\overset{x_B k_f}{\rightleftharpoons}} R^*B \qquad (1)$$

Here R is a closed (non-conducting) channel, R* an open (conducting) channel, and R*B an open channel blocked by drug molecule B (also non-conducting). The reaction $R \rightleftharpoons R^*$ represents the normal channel gating mechanism (when agonist binding is fast)[13]; β' is the effective, agonist concentration-dependent, rate constant for channel opening and α is the rate constant for shutting (approximately equal to τ_c^{-1}). The reaction $R^* \rightleftharpoons R^*B$ represents block of an open channel (with blocker concentration x_B), with association rate constant k_f, and dissociation rate constant, k_b. This mechanism predicts that, in the presence of drug, endplate current decays as the sum of rapid and slow exponential components whose rates are, respectively, speeded and slowed with increasing drug concentration. This agrees with our m.e.p.c. and noise results.

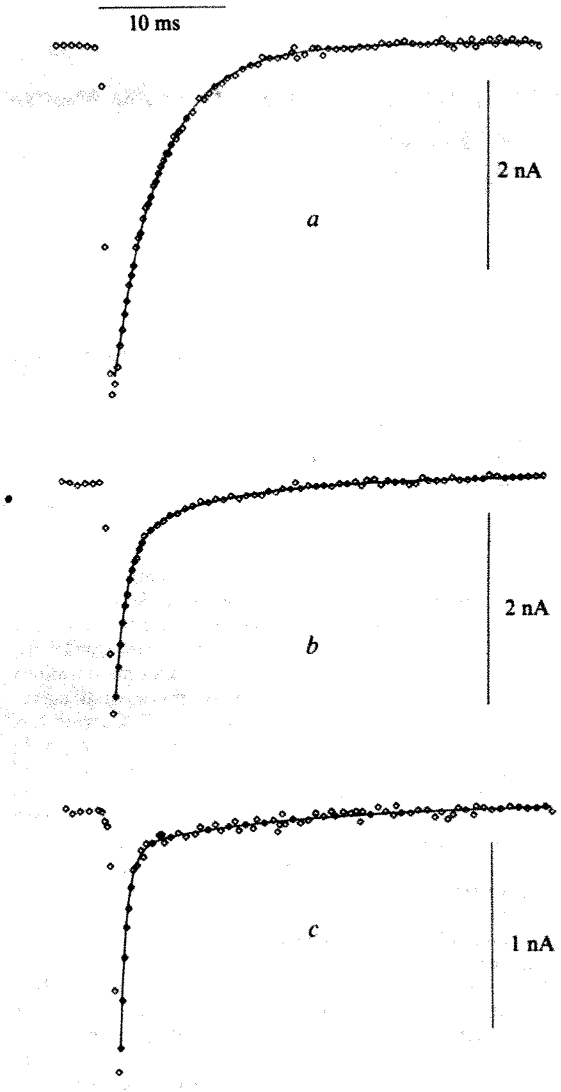


Fig. 1 Concentration dependence of the effect of benzocaine on m.e.p.c. Records are digitized, computer-averaged m.e.p.cs (8–14 individual currents) from a cutaneus pectoris muscle fibre (*Rana temporaria*) voltage-clamped to −100 mV. *a*, In the absence of benzocaine, averaged m.e.p.c. rises rapidly (within <1 ms) to a peak inward value. Current decay is fitted by a single exponential function (continuous curve) with time constant, $\tau_c = 3.8$ ms. *b*, Averaged m.e.p.c. recorded 15 min after switching to solution containing 300 μM benzocaine. Current decay fitted by the sum of two exponential components (continuous curve) with time constants, τ_f and τ_s, of 1.0 and 7.6 ms. Relative amplitude of the fast to slow exponential component at the time of the peak m.e.p.c., $I_0(f)/I_0(s) = 3.0$. *c*, Averaged m.e.p.c. obtained after 17 min in the presence of 500 μM benzocaine. The parameters of the two fitted exponentials (continuous curve) are: $\tau_f = 0.7$ ms, $\tau_s = 11.6$ ms and $I_0(f)/I_0(s) = 6.8$. Note change in current scale from *a* and *b* to *c*. Inward current is plotted in a downward direction. Muscle fibre was voltage-clamped with a conventional two-microelectrode technique and applied current measured via a KCl-agar bridge and Ag/AgCl bath electrode, connected to a virtual ground amplifier. Currents were filtered between 0.02 Hz (high pass) to 1.3 kHz (low pass). All curves were fitted using a least squares procedure. The Ringer solution contained 116.5 mM NaCl, 2.5 mM KCl, 1.5 mM $CaCl_2$ plus 50 nM tetrodotoxin and was buffered to *p*H 7.4 with 5 mM HEPES. Temperature, 11.5 °C. Muscle fibre 42-1.

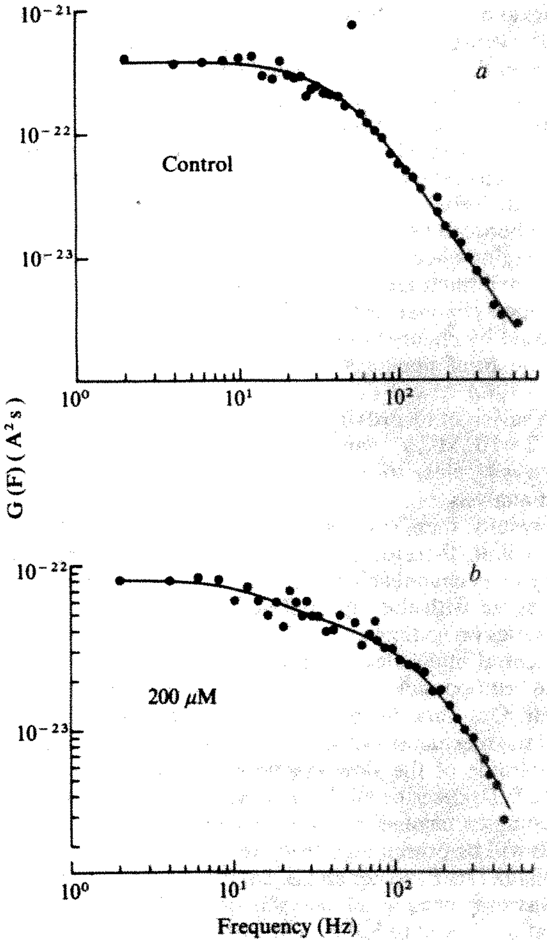


Fig. 2 Analysis of the effect of benzocaine on agonist-induced current fluctuations. *a*, Spectral density of current noise produced by bath-application of 4 μM ACh. The continuous line is a fitted single lorentzian curve; cut-off frequency, $f_c = 42.5$ Hz. *b*, Spectral density of current noise produced by application of 5 μM ACh in the presence of 200 μM benzocaine. The continuous curve is the sum of two lorentzian functions with cut-off frequencies of 17.0 and 154.8 Hz. Relative amplitude of zero frequency asymptotes of fast to slow components is 0.9. Spectra were obtained by subtracting the average of 50–60 individual spectra in the absence of agonist from the average of 50–60 spectra in the presence of agonist. Membrane potential, −70 mV; temperature, 8 °C. Cutaneus pectoris muscle fibre 43-2. Cholinesterase not inhibited.

A second prediction of the model is that the sum of the fast and slow decay rate constants (reciprocal time constants) minus the control rate constant should depend linearly on drug concentration according to; $\tau_f^{-1} + \tau_s^{-1} - \tau_c^{-1} = x_B k_f + k_b$. This prediction is tested in Fig. 3, using data pooled from many separate noise experiments. The data are fitted well by a linear relation, in agreement with the channel block model. The slope of the line yields a value for k_f of $2.6 \pm 0.3 \times 10^6\ M^{-1}\ s^{-1}$ and the intercept yields a value of $296 \pm 124\ s^{-1}$ (mean ± s.d.) for the unblocking rate, k_b. The apparent equilibrium constant, K^*, for the binding of benzocaine to its blocking site is then given by: $K^* = k_b/k_f = 114$ μM. These values are within a factor of five of those obtained for ACh channel block (at −70 mV) by charged and partially ionized drugs, QX-222 (ref. 8) and procaine[4]. The rates for channel block by benzocaine obtained in this way do not depend noticeably on membrane potential.

Recordings of single ion channel currents, illustrated in Fig. 4, show more directly the behaviour expected if benzocaine blocks open ion channels. In the absence of benzocaine (Fig. 4*a*) single channel openings give rise to rectangular current pulses. The channels stay open for a random lifetime that is exponentially distributed. In the presence of 200 μM benzocaine (Fig. 4*b*) the rectangular control current pulses are divided into bursts of rapid activity. When viewed on the expanded time scale of Fig. 4*c*, these bursts can be seen to be composed of a series of rapid openings and closings. No evidence for two channel populations, one with average open time shorter and one with average open time longer than control, was detected. Two population models have been invoked to explain biphasic m.e.p.cs in the presence of the lipophilic drug hexanol[14].

The single channel records in the presence of benzocaine are similar in appearance to records obtained in the presence of QX-222 by Neher and Steinbach[8]. Following their analysis, the bursts in benzocaine can be interpreted in the following manner: a burst begins when a channel opens ($R \rightarrow R^*$), the rapid switchings within a burst are caused by drug blocking and unblocking of the open channel ($R^* \longleftrightarrow R^*B$), until, finally, the burst is terminated by channel closure ($R^* \rightarrow R$). The mean open time, τ_0, of the brief openings within a burst should be given by $\tau_0 = (\alpha + x_B k_f)^{-1}$, and the mean blocked time, τ_b, is equal to $1/k_b$. Analysis of records like those in Fig. 4 gives values for k_f of $1.8 \pm 0.2 \times 10^6\ M^{-1}\ s^{-1}$ and for k_b of $270 \pm 20\ s^{-1}$ (mean ± s.e.m., $n = 4$), close to the rates obtained above from endplate current analysis.

Our results, then, suggest that benzocaine blocks open ACh channels and, therefore, that charge *per se* does not have a critical role in channel block by amine local anaesthetics. These results agree with the reported action of neutral amine local anaesthetics on sodium channels[1,9] and also with evidence that other neutral molecules, including the uncharged form of barbiturate compounds[12] and prednisolone[15], block endplate channels. Our data do show some deviations from the predictions of the mechanism specified in equation (1)[16]; in particular, the amplitude of the slow kinetic component is smaller than predicted. To describe all the observed effects of benzocaine, a more complex channel block mechanism must be postulated, and this will be presented elsewhere.

We thank Drs E. Neher and B. Sakmann for their advice. This work was supported by a fellowship of the Muscular Dystrophy Association (USA) to S.A.S. and by the MRC.

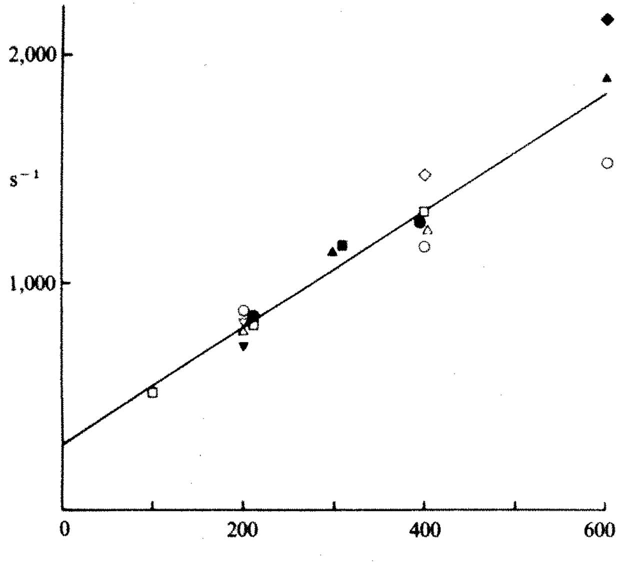


Fig. 3 Plot of $\tau_f^{-1} + \tau_s^{-1} - \tau_c^{-1}$ against benzocaine concentration. Data were obtained from noise analysis as described for Fig. 2. Each symbol represents a different experiment, all at −70 mV membrane potential. The line was fitted by the method of least squares. Temperature was between 6.5 and 10.5 °C, rates normalized to 8 °C using a Q_{10} for τ_f^{-1} of 1.7, a Q_{10} for τ_s^{-1} of 1.6 (D.C.O. and S.A.S., unpublished), and a Q_{10} for τ_c^{-1} of 2.7 (ref. 13).

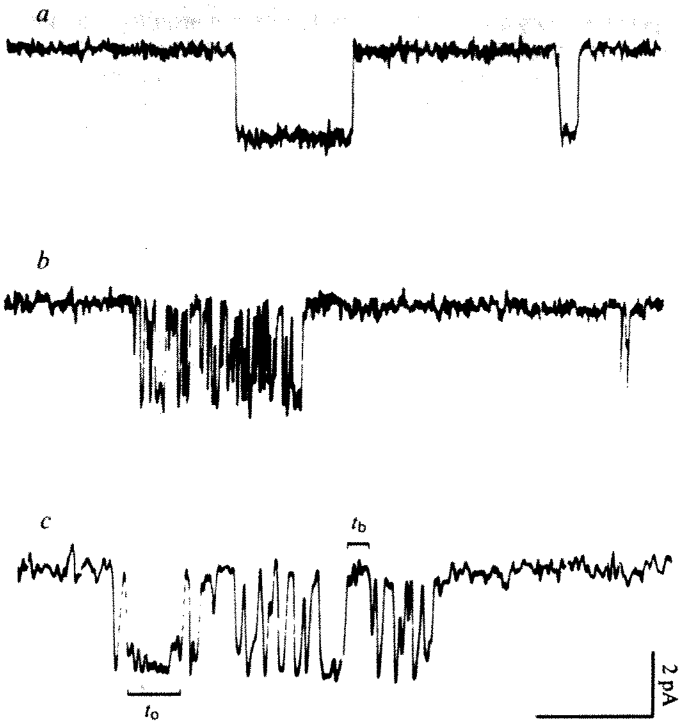


Fig. 4 Effects of benzocaine on single ACh channel currents, recorded with an extracellular patch pipette. *a*, Single channel current in the absence of benzocaine. The pipette contained Ringer solution plus 100 nM ACh. Mean channel open time, $\tau_0 = 19$ ms, was obtained by constructing a histogram of open time durations for many channel openings and fitting an exponential curve. Membrane potential, −110 mV; temperature 9 °C. Current low pass filtered at 400 Hz. Fibre 121-4. *b*, Single channel currents in the presence of 200 μM benzocaine (benzocaine in pipette as well as bath). Records were obtained from the same muscle as in *a*, but with a different fibre. Membrane potential, −130 mV; 100 nM ACh in pipette; temperature 9 °C. Current filtered at 460 Hz. Fibre 121-7. *c*, A burst in 200 μM benzocaine (from same run as *b*) at better time resolution. Current filtered at 800 Hz. The brackets indicate one open interval (t_0) and one blocked interval (t_b). Histograms of open and blocked times were well fitted by single exponentials, yielding a mean open time of 2.8 ms and a mean blocked time of 3.5 ms. Time calibration bar corresponds to 200 ms in panels *a*, *b* and 50 ms in panel *c*. Downward deflection indicates inward current. Recordings obtained from the extrajunctional membrane of 4–6-week denervated semitendinosus muscle fibres. Single channel currents were recorded using methods described by Neher *et al.*[17] and Neher[18]. Suction was applied to the pipette interior to increase the seal resistance between pipette and fibre membrane to values greater than $5 \times 10^8\ \Omega$ (see ref. 18).

Received 11 September; accepted 4 December 1980.

1. Hille, B. *J. gen. Physiol.* **69**, 497–515 (1977).
2. Courtney, K. *J. Pharmac. exp. Ther.* **195**, 225–236 (1975).
3. Strichartz, G. R. *J. gen. Physiol.* **62**, 37–57 (1973).
4. Adams, P. R. *J. Physiol., Lond.* **268**, 291–318 (1977).
5. Ruff, R. L. *J. Physiol., Lond.* **264**, 89–124 (1977).
6. Marty, A. *J. Physiol., Lond.* **278**, 237–250 (1978).
7. Lester, H. A., Krouse, M. E., Nass, M. M., Wasserman, N. H. & Erlanger, B. F. *Nature* **280**, 509–510 (1979).
8. Neher, E. & Steinbach, J. H. *J. Physiol., Lond.* **277**, 153–176 (1978).
9. Rimmel, C., Walle, A., Kessler, H. & Ulbricht, W. *Pflügers Arch. ges. Physiol.* **376**, 105–118 (1978).
10. Steinbach, A. B. *J. gen. Physiol.* **52**, 144–161 (1968).
11. Katz, B. & Miledi, R. *J. Physiol., Lond.* **249**, 269–284 (1975).
12. Adams, P. R. *J. Physiol., Lond.* **260**, 531–552 (1976).
13. Magleby, K. L. & Stevens, C. F. *J. Physiol., Lond.* **223**, 173–197 (1972).
14. Gage, P. W., McBurney, R. N. & Schneider, G. T. *J. Physiol., Lond.* **244**, 409–429 (1975).
15. Dreyer, F., Peper, K., Sterz, R., Bradley, R. J. & Müller, K. D. *Prog. Brain Res.* **49**, 213–223 (1979).
16. Ogden, D. C. & Siegelbaum, S. A. *Br. J. Pharmac.* **69**, 336P–337P (1980).
17. Neher, E., Sakmann, B. & Steinbach, J. H. *Pflügers Arch. ges. Physiol.* **375**, 219–228 (1978).
18. Neher, E. in *Techniques in Cellular Physiology* (ed. Baker, P. F.) (Elsevier, North Holland, in the press).

Alterations in brain cholecystokinin receptors after fasting

A. Saito*†, J. A. Williams*† & I. D. Goldfine*†‡

* Cell Biology Laboratory and Department of Medicine, Mount Zion Hospital and Medical Center, San Francisco, California 94120
Departments of † Physiology and ‡ Medicine, University of California, San Francisco, California 94143

Cholecystokinin (CCK) is the parent molecule of a family of polypeptide hormones, some of which are secreted from the small intestine after food ingestion and stimulate both exocrine secretion and gall bladder contraction[1–4]. CCK-like molecules have also been identified in the central and peripheral nervous systems[5–12]. The significance of CCK in the brain is unknown, but it could mediate satiety. The satiety produced by introduction of food into the intestine[13] can be mimicked by systemic injections of CCK and its analogues[14–17]—these hormones are also effective when injected into the hypothalamus[18] and the cerebral ventricle[19]. Straus and Yalow showed that brain CCK levels were reduced in the cerebral cortex of both genetically obese (*ob/ob*)[20] and normal mice after a 2–5-day fast[21]. However, Schneider[22] detected no such reduction in similar conditions and other studies have suggested a peripheral rather than central action on satiety[23,24]. Using CCK of high specific activity, specific CCK receptors have been measured both in pancreatic acini and in various brain regions[25–28]. We show here that fasting in mice significantly increases CCK binding due to an increased number of CCK receptors in the olfactory bulb and hypothalamus, but not in other brain regions. In contrast, insulin binding to its receptors was not altered by fasting. As the hypothalamus is known to regulate appetite[29,30], this finding supports the concept that CCK regulates satiety through interaction with this brain region.

Male white Swiss mice were either allowed to feed *ad libitum* or underwent a 42-h fast, but were allowed water. In fasted animals, body weight fell from 33.8±0.45 to 30.8±0.89 g ($n = 16$). Both fed and fasted animals were then rapidly decapitated, the brain removed, specific regions isolated and brain particles prepared[26]. Specific ^{125}I-BH-CCK (CCK conjugated to labelled Bolton-Hunter reagent) binding to receptors was then measured[26] (Table 1). In both fed and fasted rats, specific CCK binding varied in different regions of the brain. However, when fed and fasted animals were compared, a significant increase in binding was seen in the olfactory bulb and hypothalamus of fasted animals (Table 1). Although several other brain regions, including cerebral cortex, hippocampus, midbrain, caudate and hindbrain, tended to show an increase in binding in fasted animals, the changes were insignificant (Table 1).

Scatchard analyses of CCK binding were then carried out in the olfactory bulb and hypothalamus; the cerebral cortex was studied concurrently as a control. Fasting was accompanied by a significant increase in the CCK binding capacity of both the olfactory bulb and hypothalamus, but not the cerebral cortex (Fig. 1). Preliminary studies indicate that 48 h after refeeding of fasted mice, the number of CCK receptors returns to the value seen in fed mice. In all brain areas studied, the binding affinity (K_d) was ~1 nM, a value similar to that previously reported for rat cerebral cortex[26].

In contrast to changes in CCK binding, no difference in CCK degradation by brain particles was detected in any of the brain regions from fasted animals. As insulin is another polypeptide hormone that binds to specific receptors in various regions of the brain[31] and these receptors do not change in various metabolic states[32,33], we used insulin binding to its receptors as a control. No differences in insulin binding to its receptors in various brain regions, including hypothalamus and olfactory bulb, were detected between fed and fasted mice.

The study indicates that fasting increases the number of CCK receptors in specific regions of mouse brain. Our data and those in the literature therefore support the view that metabolic alterations may change not only brain CCK levels, but also CCK receptor concentrations. There is evidence in rats and other rodents that the hypothalamus mediates satiety[29,30]. Bilateral destruction of the ventromedial hypothalamus of rats results in an increase in both food intake and body weight[29,30]. Furthermore, in rats with lesions in the ventromedial hypothalamus, there is reduced suppression of feeding by caerulein (a CCK analogue)[18]. The increase in the number of receptors for CCK in the hypothalamus after fasting shown here suggests that alterations in these receptors may be important in hunger and satiety.

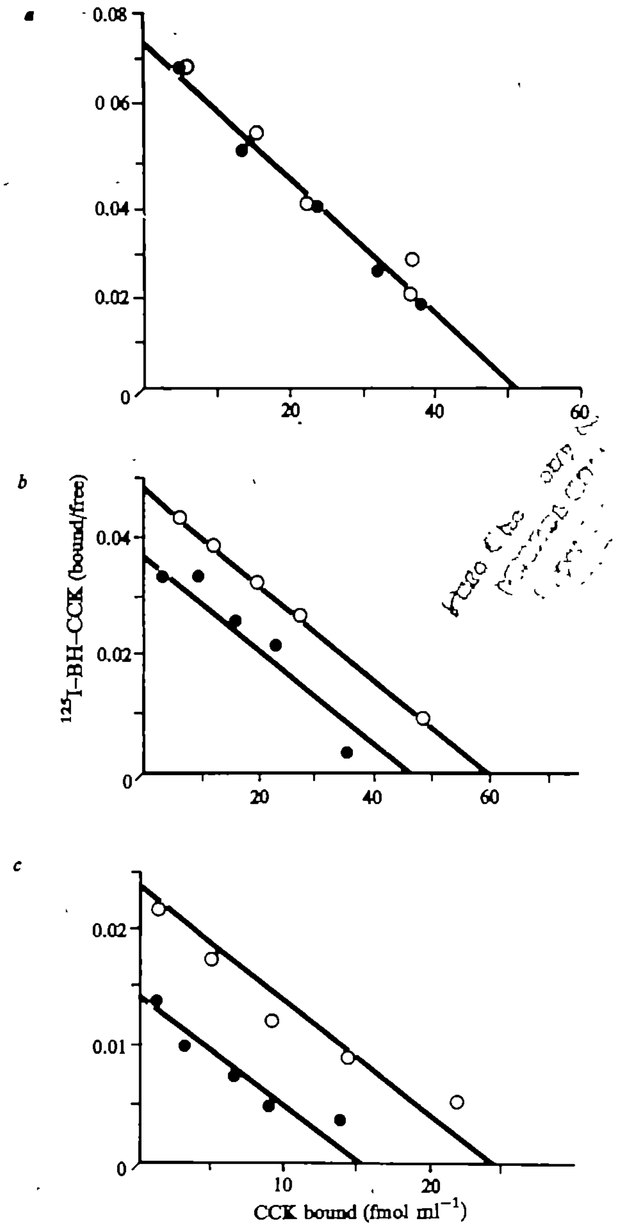


Fig. 1 Scatchard plots of specific ^{125}I-BH-CCK binding to cerebral cortex (*a*), olfactory lobe (*b*) and hypothalamus (*c*). ●, Fed; ○, fasted. Binding, expressed per mg protein per ml, was carried out as specified in Table 1. The ligand was added at 15 pM and unlabelled CCK was added in concentrations from 10 pM to 100 nM.

0028-0836/81/070599–02$01.00

Table 1 Effect of fasting on CCK receptors in mouse brain

Region		Specific CCK binding (fmol per mg protein × 10^2)		% Increase in fasted
		Fed	Fasted	
Olfactory bulb	(11)	195 ± 22	285 ± 30	46+
Hypothalamus	(9)	99 ± 12	141 ± 13	42+
Hindbrain	(7)	70 ± 19	90 ± 11	29
Hippocampus	(7)	118 ± 10	139 ± 19	18
Cerebral cortex	(8)	397 ± 36	469 ± 41	18
Midbrain	(7)	103 ± 13	112 ± 11	9
Caudate	(7)	96 ± 8	92 ± 14	0

Mouse brain particles were prepared by a modification of the method of Uhl *et al.*[34]. Male Swiss mice (30–35 g), either fed or fasted for 42 h, were kept at 24 °C on a light–dark cycle changing at 9 a.m. and 9 p.m. The animals were decapitated between 9:30 and 11:00 a.m., the brain dissected on ice and the various regions homogenized in 10 volumes of 50 mM Tris-HCl (*p*H 7.4) at 4 °C by the use of a motor-driven Teflon-glass homogenizer[26]. The homogenate was centrifuged at 42,000*g* for 15 min at 4°C, resuspended in an equal volume of buffer and similarly recentrifuged. The pellet was then resuspended by homogenization in buffer containing 180 mM NaCl, 4.7 mM KCl, 5 mM $MgCl_2$, 1 mM EGTA, 1 mg ml^{-1} bacitracin, 5 mg ml^{-1} bovine serum albumin and 10 mM HEPES (*p*H 7.4)[26]. To measure binding, 50 pM ^{125}I-BH-CCK[25,26] was added in triplicate to plastic microcentrifuge tubes in a 300-µl final volume along with brain particles (final protein concentration 0.3–0.5 mg ml^{-1}) at 24 °C. After 120 min, the tubes were centrifuged at 10,000*g* in a Beckman 152 microcentrifuge at 24 °C. The pellets were rinsed with cold buffer, recentrifuged, the tips of the tubes cut off and radioactivity determined. Nonspecific binding was determined by adding unlabelled 10^{-6}M CCK_8 with the labelled hormone and this value was subtracted from total binding to yield specific binding[26]. Each value is the mean ± s.e. The number of experiments performed is given in parentheses. +, $P < 0.05$ (fasted compared with fed). In every experiment, CCK binding to its receptors in olfactory bulb and hypothalamus was greater in fasted than fed mice.

This research was supported by NIH grant AM26422 and the Elise Stern Haas Research Fund, Harold Brunn Institute, Mount Zion Hospital and Medical Center. We thank Dr H. Sankaran for preparing the ^{125}I-BH-CCK.

Received 1 August; accepted 12 November 1980.

1. Pearse, A. G. E. *Nature* **262,** 92–94 (1976).
2. Dockray, G. J. *Fedn Proc.* **38,** 2295–2301 (1979).
3. Jorpes, J. E. & Mutt, J. *Secretin, Cholecystokinin, Pancreozymin, and Gastrin* (eds Jorpes, J. E. & Mutt, J.) 1–144 (Springer, Berlin, 1973).
4. Dockray, G. J. *Nature* **270,** 359–361 (1977).
5. Rehfeld, J. F. *J. biol. Chem.* **253,** 4016–4021 (1978).
6. Vanderhaeghen, J. J., Signeau, J. C. & Gepts, V. *Nature* **257,** 604–605 (1975).
7. Dockray, G. J., Gregory, R. A., Hutchison, J. B., Harris, J. I. & Runswick, M. *Nature* **274,** 711–713 (1978).
8. Larsson, L.-I. & Rehfeld, J. F. *Brain Res.* **165,** 201–218 (1979).
9. Straus, E., Muller, J. E., Choi, H.-S., Paronetto, F. & Yalow, R. S. *Proc. natn. Acad. Sci. U.S.A.* **74,** 3033–3034 (1977).
10. Innis, R. B., Correa, F. M. A., Uhl, G. R., Schneider, B. & Snyder, S. H. *Proc. natn. Acad. Sci.* U.S.A. **76,** 521–525 (1979).
11. Loren, I., Alumets, J., Hakanson, R. & Sundler, F. *Histochemistry* **59,** 249–257 (1979).
12. Vanderhaeghen, J. J., Lotstra, F., De Mey, J. & Gilles, C. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1190–1194 (1980).
13. Smith, G. P., Gibbs, J. & Young, R. C. *Fedn Proc.* **33,** 1146–1149 (1974).
14. Gibbs, J., Young, R. C. & Smith, G. P. *J. comp. physiol. Psychol.* **84,** 488–495 (1973).
15. Gibbs, J., Young, R. C. & Smith, G. P. *Nature* **245,** 323–325 (1973).
16. Houpt, T. R., Anika, S. M. & Wolff, N. C. *Am. J. physiol.* **235,** R23–R28 (1978).
17. Stacher, G., Bauer, H. & Steinringer, H. *Physiol. Behav.* **23,** 325–331 (1979).
18. Stern, J. J., Cudillo, C. A. & Kruger, J. *J. comp. physiol. Psychol.* **90,** 484–490 (1976).
19. Della-Fera, M. A. & Baile, C. A. *Science* **206,** 471–473 (1979).
20. Straus, E. & Yalow, R. S. *Science* **203,** 68–69 (1978).
21. Straus, E. & Yalow, R. S. *Life Sci.* **26,** 969–970 (1980).
22. Schneider, B. S., Monahan, J. W. & Hirsch, J. *J. clin. Invest.* **64,** 1348–1356 (1979).
23. Nemeroff, C. B. *et al. Science* **200,** 793–794 (1978).
24. Smith, G. P. & Cushin, B. J. *Neurosci. Abstr.* **4,** 180 (1978).
25. Sankaran, H., Deveney, C. W., Goldfine, I. D. & Williams, J. A. *J. biol. Chem.* **254,** 9349–9351 (1979).
26. Saito, A., Sankaran, H., Goldfine, I. D. & Williams, J. A. *Science* **208,** 1155–1156 (1980).
27. Jensen, R. T., Lemp, G. F. & Gardner, J. D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2079–2083 (1980).
28. Innis, R. B. & Snyder, S. H. *Eur. J. Pharmac.* **65,** 123–124 (1980).
29. Kennedy, G. C. *Proc. R. Soc.* **B137,** 535–549 (1950).
30. Mayer, J. & Thomas, D. W. *Science* **156,** 328–337 (1967).
31. Havrankova, J., Brownstein, M. & Roth, J. *Nature* **272,** 827–829 (1978).
32. Pacold, S. T. & Blackard, W. G. *Endocrinology* **105,** 1452–1457 (1979).
33. Havrankova, J., Roth, J. & Brownstein, M. J. *J. clin. Invest.* **64,** 636–642 (1979).
34. Uhl, G. R., Bennet, J. P. Jr & Snyder, S. H. *Brain Res.* **130,** 299–313 (1977).

Vitamin D-dependent phosphorylation of an intestinal protein

P. W. Wilson & D. E. M. Lawson

Dunn Nutritional Laboratory, University of Cambridge and Medical Research Council, Milton Road, Cambridge CB4 1JX, UK

The complete pattern of response of target tissues to steroid hormones is still unclear although it seems that the initial biochemical events involve the transcriptional control of protein synthesis. A still outstanding problem concerns the role of these hormone-dependent proteins in the physiological response. Synthesis of several intestinal proteins is known to be stimulated by 1,25-dihydroxyvitamin D (1,25-$(OH)_2D$)[1]; at least one, calcium-binding protein (CaBP) by a transcriptional event[2]. Steroid hormones also stimulate protein phosphorylation[3] and we show here that 1,25-$(OH)_2D$ acting *in vivo* stimulates the ATP phosphorylation of a chick intestinal brush-border membrane protein of molecular weight (MW) 84,000. The phosphate is attached to the protein through a serine or threonine residue. The reaction is stimulated by Ca^{2+} but does not require Mg^{2+}, and kinetic studies showed that 1,25-$(OH)_2D$ acts on this process by increasing the number of sites rather than the affinity of the substrate for the phosphate group. This phosphorylated protein has similar characteristics to those of the intestinal membrane protein whose synthesis is stimulated by 1,25-$(OH)_2D$.

The realization that more intestinal components than CaBP are involved in vitamin D-stimulated Ca absorption[4–6] gave fresh impetus to the search for their identification. 1,25-$(OH)_2D$ stimulates the synthesis of two brush-border proteins[7,8], in addition to the cytoplasmic CaBP[9] and a recently described mitochondrial protein[10]. One of these brush-border proteins is actin[11] and the other has a MW of 84,000. More evidence is required before the latter protein and the one undergoing ATP phosphorylation can be regarded as identical but it seems unlikely that there are two proteins of very similar molecular size in the intestinal brush border which are affected by vitamin D. We refer to the protein which undergoes phosphorylation as vitamin D-responsive protein (DRP). There have been preliminary reports of 1,25-$(OH)_2D$-dependent phosphorylation of intestinal brush-border membranes[12,13].

Following the incubation of chick intestinal brush-border membranes in a buffer containing [γ-^{32}P]ATP and Mg^{2+} for short periods of time, many of the protein components become phosphorylated (Fig. 1*a*) but in the absence of Mg^{2+} only one protein (MW 84,000) is phosphorylated (Fig. 1*b*). The phosphorylation reaction was present in intestinal brush borders or the membranes and core fractions derived from them. Treatment of the phosphorylated DRP with hydroxylamine did not release the phosphate from the protein (Fig. 1*c*) indicating that the product of this reaction is not an acyl phosphate of the type found in ATPases[14]. Furthermore the phosphorylated protein was stable to 5% trichloroacetic acid at 95 °C for 5 min but was hydrolysed by 1 M NaOH in the same conditions, which suggests that the phosphate acceptor in the protein is serine or threonine[15]. In incubations containing the brush-border fraction the rate of formation of phosphorylated DRP increased for up to 10 min and then remained substantially unchanged for a further 10 min (Fig. 2). However, in the case of the brush-border membrane, DRP phosphorylation was complete within 10 s, with significant dephosphorylation occurring over the next 10 min (Fig. 2). The effect of varying ATP levels on the initial phosphorylation rate of the brush border and its component

0028-0836/81/070600-03$01.00

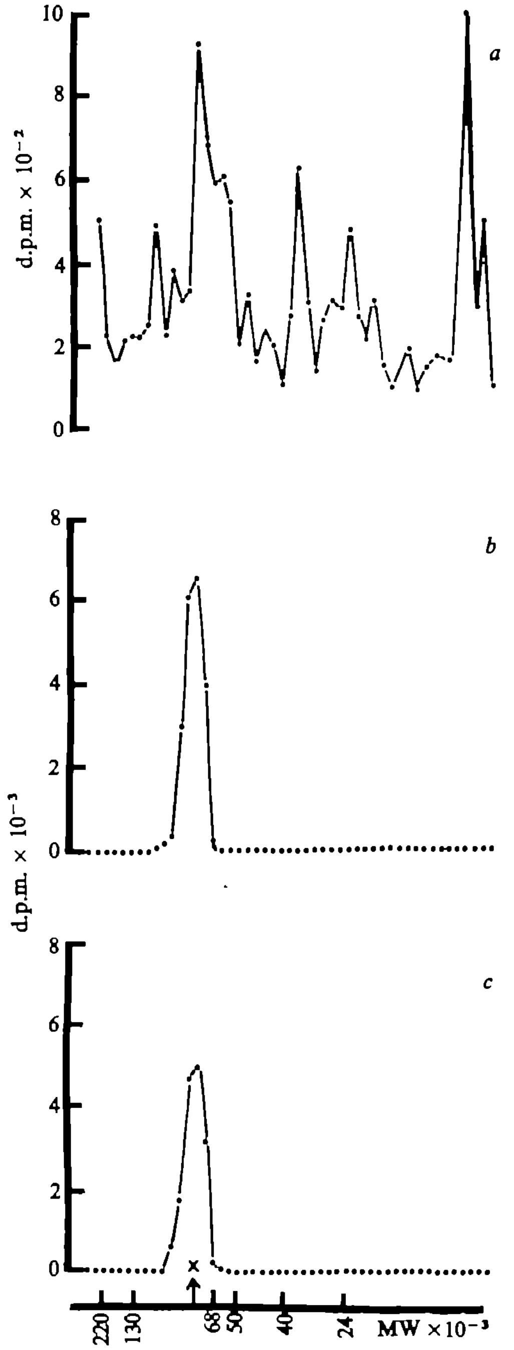


Fig. 1 Chick intestinal brush borders were prepared[8], washed twice in 40 mM Tris-HCl, *p*H 7.4, and resuspended in this buffer for storage in liquid N_2 before use. *a*, The brush borders (100 μg) were incubated at 0 °C in a medium (50 μl) of 40 mM Tris-HCl, *p*H 7.4, 2.5 mM Mg^{2+} and 2.5 μCi [γ-^{32}P]ATP for 30 s. The reaction was stopped with a dissociating solution (50 μl) to give a final concentration of 12 mg Tris, 15 mg dithiothreitol, 10 mg SDS and 100 μl glycerol ml^{-1} by heating at 70 °C for 30 min. *b*, The brush borders (100 μg) were incubated at 0 °C in a medium (100 μl) comprising: 5 mM theophylline, 10 mM NaF, 1 mM quercetin, 1 mM ouabain, 2.5 μCi[γ-^{32}P]ATP and 40 mM Tris-HCl, *p*H 7.4. The reaction was stopped with 1 ml 5% trichloroacetic acid (TCA), 2 mM H_3PO_4. The precipitate was dissolved in 100 μl of the dissociating solution to give the final concentrations as above. *c*, Brush borders incubated as in *b*, but the TCA precipitate was treated with 100 μl of 5 M hydroxylamine, *p*H 5.2, at 25 °C for 15 min before addition of 25 μl of dissociating solution (to give the appropriate concentration as above) and heating at 70 °C for 30 min. The proteins were fractionated by discontinuous gel electrophoresis[11]. The gels were sliced in 1-mm sections and ^{32}P measured in a liquid scintillation counter. ×, Position of 84,000-MW protein[7].

Table 1 Kinetic characteristics of the phosphorylation reaction undergone by intestinal brush border, and its membrane and core components

Treatment	Ca^{2+}	Brush border		Membrane		Core	
		V_{max}	K_m	V_{max}	K_m	V_{max}	K_m
−D	−	0.09	0.51	0.88	1.01	0.98	0.26
−D	+	0.21	1.23				
+D	−	0.60	0.47	3.49	0.78	2.04	0.27
+D	+	1.03	1.16	3.03	1.97	1.97	0.77
+1,25-$(OH)_2D_3$	−	0.25	0.49	1.76	1.09	1.11	0.27
+1,25-$(OH)_2D_3$	+	0.48	1.09				

Intestinal brush borders and their membrane and core components were prepared as described in the legends to Figs 1, 2. Incubations were carried out and stopped as described in Fig. 2 legend except that 0.25 mM Ca^{2+} was added to the incubation medium where indicated. The concentrations of ATP used were in the range 0.2–10 μM ATP. V_{max} (pmol per 10 μg per 15 s) and K_m (μM) values for the three fractions were obtained from a Lineweaver–Burke analysis of the initial velocities measured from 0–15 s. −D, vitamin D-deficient birds; +D, vitamin D-deficient birds treated with 12.5 μg of vitamin D_3 for 72 h before killing; +1,25-$(OH)_2D_3$, vitamin D-deficient birds treated with 125 ng of 1,25-$(OH)_2D_3$ for 4 h before killing.

parts is shown in Table 1. Greater phosphorylating activity was detected in the two components than in the whole brush border. The V_{max} of the reaction using the membrane and core fractions was almost sixfold and threefold, greater, respectively, than the V_{max} achieved by the brush borders.

A notable feature of this phosphorylation reaction is the lack of a requirement for Mg^{2+}. Brush borders from the intestine of vitamin D-replete chicks incubated in the usual medium continued to phosphorylate the DRP at an unchanged rate in the presence of EDTA or *trans*-1,2-diaminocyclohexanetetraacetic acid, a specific chelator of Mg^{2+}. However, the data in Table 1 show that Ca^{2+} stimulates the initial phosphorylation rate of the brush-border fraction but not of the derived membranes. The V_{max} of the reaction using the brush-border fraction was increased by about 70% and the K_m increased from 0.47 μM to 1.16 μM. In the case of the membrane fraction, the addition of Ca^{2+} had no effect on the V_{max} of the reaction but increased the K_m from 0.78 to 1.97 μM. Incubations carried out with brush borders, their membranes or their cores, from the intestine of vitamin D-deficient chicks showed that the initial DRP phosphorylation rate was significantly reduced compared with the activity shown by the same cell fraction from either vitamin D-replete birds or deficient chicks dosed with 1,25-$(OH)_2D_3$ (Table 1). This decrease in phosphorylation rate was reflected in a reduced V_{max} for the reaction, but the K_m for the reaction was virtually unchanged by the vitamin D status of the birds. An increase in the initial phosphorylation rate of DRP is observed within 4 h of a dose of 1,25-$(OH)_2D_3$ to rachitic chicks.

ATPases, components of ion pumps in membranes, are very rapidly phosphorylated with the formation of an acylphosphate intermediate[14]. Although the very rapid rate of phosphorylation occurring with the membrane fraction is typical of that found with ATPase enzymes, the evidence given above of the nature of the phosphate acceptors suggests a kinase type of phosphorylation reaction. If this is the case, the kinase is not of the class stimulated by cyclic nucleotides. The different rates of DRP phosphorylation using either brush borders or the membranes raises some intriguing questions concerning the control of the reaction and whether reaction rates are governed by the accessibility of the enzyme to the substrate. The greater level of activity shown by both the membranes and core fractions compared to the brush borders from which they were derived may indicate that the enzyme is on the inside of the brush-border membrane, or that there are factors[13] controlling the phosphorylation reaction which are released in formation of the membrane fraction so that the reaction proceeds very rapidly to

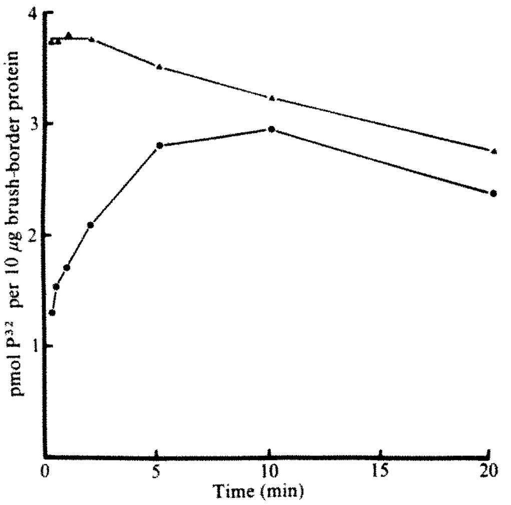


Fig. 2 Membranes of the chick intestinal brush borders were prepared by treating 2 mg of brush-border protein with 100 μl of 1% Triton X-100 in 40 mM Tris-HCl, *p*H 7.4, for 16 h at 4 °C. The solution was then centrifuged at 10,000*g* for 30 min—a pellet of the brush-border core material was produced, leaving the membranes suspended in the supernatant. The core material was resuspended in 100 μl of 1% Triton X-100 in 40 mM Tris-HCl, *p*H 7.4. The brush borders (5 μg of protein), or the membranes from 5 μg of brush-border protein, were incubated in the medium (50 μl) described in Fig. 1*b* legend plus 10 μM ATP, for the times indicated. the reactions were stopped with a solution (1.0 ml) of 5% TCA, 2 mM H_3PO_4 followed by addition of 500 μg of albumin as carrier. The precipitate was collected by centrifugation at 3,000 r.p.m. for 10 min, dissolved in 100 μl of 1 M NaOH at 0 °C and the protein re-precipitated with the above solution[16]. This precipitate was collected on Whatman glass fibre filters (GF/C), washed with 20 ml of ice-cold 10% TCA and the filters dried and ^{32}P measured with a scintillation counter. All assays were carried out in triplicate.

completion. Determination of the role of the DRP and its phosphorylation in vitamin D-stimulated calcium absorption must await further information but observations on the turnover of the phosphate group on phosphorylated DRP show that this is too slow for the phosphate to be involved in Ca absorption in a stoichiometric manner. If DRP is involved in Ca absorption, it is likely that the appropriate conformation requires the protein to be phosphorylated. These results indicate for the first time that ATP may be important in the response of the intestine to 1,25-$(OH)_2D$ but that the proportion of total cellular ATP involved is small.

Received 30 July; accepted 2 December 1980.

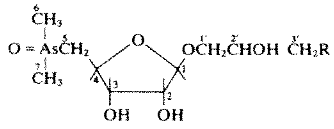
1. Lawson, D. E. M. in *Vitamin D* (ed. Lawson, D. E. M.) 167–200 (Academic, London, 1978).
2. Spencer, R., Charman, M. & Lawson, D. E. M. *Biochem. J.* **175,** 1089–1094 (1978).
3. Greengard, P. *Fedn Proc.* **38,** 2208–2217 (1979).
4. Spencer, R., Charman, M., Wilson, P. & Lawson, D. E. M. *Nature* **263,** 161–163 (1975).
5. Morrissey, R. L., Zolock, D. T., Bikle, D. D., Empson, R. N. & Bucci, T. J. *Biochim. biophys. Acta* **538,** 23–33 (1978).
6. Thomasset, M., Cuisinier-Gleizes, P. & Mathieu, H. *FEBS Lett.* **107,** 91–94 (1979).
7. Wilson, P. W. & Lawson, D. E. M. *Biochim. biophys. Acta* **497,** 805–811 (1977).
8. Max, E. E., Goodman, D. B. P. & Rasmussen, H. *Biochim. biophys. Acta* **511,** 224–239 (1978).
9. Wasserman, R. H., Fullmer, C. S. & Taylor, A. N. in *Vitamin D* (ed. Lawson, D. E. M.) 133–166 (Academic, London, 1978).
10. Hobden, A. N., Harding, M. & Lawson, D. E. M. *Nature* **288,** 718–720 (1980).
11. Wilson, P. W. & Lawson, D. E. M. *Biochem. J.* **173,** 627–631 (1978).
12. Lawson, D. E. M., Harding, M., Lane, S. & Wilson, P. W. in *Vitamin D, Basic Research and its Clinical Application* (ed. Norman, A. W.) 711–718 (de Gruyter, Berlin, 1979).
13. Wasserman, R. H. & Brindak, M. E. in *Vitamin D, Basic Research and its Clinical Application* (ed. Norman, A. W.) 703–709 (de Gruyter, Berlin, 1979).
14. Skou, J. C. *Curr. Topics Bioenerg.* **4,** 357 (1971).
15. Weller, M. in *Protein Phosphorylation,* 3 (Pion, London, 1979).
16. Rubin, C. S., Bechtel, P. J. & Krebs, E. G. *Meth. Enzym.* **38,** 308–315 (1975).

Arseno-sugars from brown kelp (*Ecklonia radiata*) as intermediates in cycling of arsenic in a marine ecosystem

J. S. Edmonds & K. A. Francesconi

Western Australian Marine Research Laboratories, PO Box 20,
North Beach, Western Australia 6020

The high concentration (relative to seawater) of arsenic in many marine animals eaten as human food has stimulated interest in the cycling of arsenic in the marine environment[1]. Although arsenic is present as arsenobetaine [$(CH_3)_3As^+CH_2COO^-$] in the western rock lobster[2] (*Panulirus cygnus*), the dusky shark[3] (*Carcharhinus obscurus*) and the school whiting[4] (*Sillago bassensis*) it is not clear what intermediate stages are involved in the biosynthesis of this compound from arsenate, the major form of arsenic in seawater[5]. We now report the isolation of the two main arsenical constituents of the brown kelp, *Ecklonia radiata*, and their identification as a 2-hydroxy-3-sulphopropyl-5-deoxy-5-(dimethylarsenoso)furanoside and a 2,3-dihydroxypropyl-5-deoxy-5-(dimethylarsenoso)furanoside. A *β*-ribo structure for the sugar system is strongly indicated in each case (*A* and *B* below).



A, R = $-SO_3H$
B, R = $-OH$

***Ecklonia* is the major organism that concentrates arsenic in the coastal ecosystem to which the western rock lobster and school whiting belong. It is clear that the compounds described here could readily be further metabolized to arsenobetaine and may well be the source of arsenobetaine in marine fauna associated with the region.**

Two water-soluble arsenic compounds (designated compounds *A* and *B*), which together accounted for 81% of the arsenic present in *Ecklonia*, were extracted from freshly collected alga (11 kg) with methanol. Compounds *A* and *B* were purified by chromatography on Sephadex LH-20 and ion-exchange Sephadex resins, and by HPLC and preparative layer chromatography. 30 mg of compound *A* and 5 mg of compound *B* were obtained as colourless syrups. Compound *A* was acidic and contained sulphur whereas compound *B* lacked sulphur and showed neither strongly acidic nor strongly basic character. Compound *A* contained: C, 26.9%, H, 5.3%; As, 17.5%; S, 7.5%. $C_{10}H_{21}O_9AsS$.$3H_2O$ requires: C, 26.9%; H, 6.0%; As, 16.8%; S, 7.2%. Compound *B* contained: As, 20.0%. $C_{10}H_{20}O_7As.3H_2O$ requires: As, 19.6%. The IR spectrum of compound *A* showed a strong absorbance at 845 cm^{-1} indicating the presence of an arsine oxide[6]. The IR spectra of *β*-ribose and 2,3-dihydroxy-1-propane sulphonic acid lacked absorbances in this region. The structures of compounds *A* and *B* were determined chiefly by NMR considerations.

The 270-MHz ^1H NMR spectra of compounds *A* and *B* (Table 1) differed substantially in only two respects. First, the eight-line (AB part of an ABX system) centred at δ 3.09 in the spectrum of compound *A* was replaced by a similar eight-line system centred at δ 3.63 in the spectrum of *B*; second, the three-proton multiplet at δ 4.25 in *A* became a two-proton multiplet (centred at δ 4.27) in *B*, with the appearance of a one-proton multiplet at δ 3.91. Irradiation at δ 4.27 (compound *A*) resulted in the collapse to AB quartets of all three eight-line (AB parts of ABX systems) systems (δ 2.59, 3.09, 3.71). Thus, because of the coincidence of methine signals at δ 4.27, limited structural information was obtained. However, irradiation at δ

0028-0836/81/070602-03$01.00

Table 1 270 MHz ^1H NMR data for compounds *A* and *B* and ^1H NMR data for model compounds

		Compound *A*				Compound *B*			
Protons*	Positition	Chemical shift δ p.p.m. relative to $(CH_3)_4$ Si	Geminal coupling constants	Vicinal coupling constants	Results† of spin decoupling at *	Chemical shift δ p.p.m. relative to $(CH_3)_4$ Si	Geminal coupling constants	Vicinal coupling constants	Results† of spin decoupling at *
Two CH_3 groups	6, 7	1.89, 1.86				1.88, 1.86			
CH	1	5.02		0		5.04		0	
CH	2	4.13‡		4.0 Hz		4.15‡		3.8 Hz	
CH	3	Multiplet centred at 4.25‡		§	*	Multiplet centred at 4.27‡		§	*
CH	4	Multiplet centred at 4.25		§	*	Multiplet centred at 4.27		§	*
$H_A-\overset{\vert}{\underset{\vert}{C}}-H_B$	5	2.52, 2.65	13.6 Hz	10.1, 3.5 Hz	AB quartet $J = 13.6$ Hz $\Delta\delta AB = 0.13$ p.p.m.	2.50, 2.66	13.6 Hz	9.9, 3.2 Hz	AB quartet $J = 13.6$ Hz $\Delta\delta AB = 0.16$ p.p.m.
$H_A-\overset{\vert}{\underset{\vert}{C}}-H_B$	1′	3.63, 3.79	10.4 Hz	3.5, 5.5 Hz	AB quartet $J = 10.4$ Hz $\Delta\delta AB = 0.16$ p.p.m.	3.62, 3.78	10.5 Hz	2.7, 6.2 Hz	AB quartet $J = 10.5$ Hz $\Delta\delta AB = 0.16$ p.p.m.
CH	2′	Multiplet centred at 4.25		§	*	Multiplet centred at 3.91		§	*
$H_A-\overset{\vert}{\underset{\vert}{C}}-H_B$	3′	3.03, 3.14	14.6 Hz	6.8, 5.7 Hz	AB quartet $J = 14.6$ Hz $\Delta\delta AB = 0.11$ p.p.m.	3.59, 3.66	11.6 Hz	6.2, 4.9 Hz	AB quartet $J = 11.6$ Hz $\Delta\delta AB = 0.07$ p.p.m.

Model compounds (**Table 1 continued**)

Structure with relevant protons indicated (*italic*)	Chemical shift δ p.p.m. relative to $(CH_3)_4Si$	Coupling constants
$(\mathit{CH_3})_3\ As^+\ CH_2COO^-$ $(\mathit{CH_3})_4As^+\ I^-$	1.87 (ref. 2) 1.87–1.95 (ref. 13)	
anomeric protons in sugar systems	4.5–5.5 (ref. 14)	‖
—*CH* OH in sugar systems	3.0–4.5 (ref. 14)	
$HO\ CH_2\ \mathit{CH_2}\ AsO_3\ H_2$	Triplet centred at 2.55	7.5 Hz
Compound *A*: $\mathit{CH_2}$ OH CHOHCH$_2$SO$_3$H	3.58	¶
Compound *B*: $\mathit{CH_2}$ OH CHOH $\mathit{CH_2}$ OH	3.57	¶
Compound *A*: CH_2 OH *CH*OH CH_2 SO_3H	4.07	¶
Compound *B*: CH_2 OH *CH*OH CH_2OH	3.68	¶
Compound *A*: CH_2OH CHOH $\mathit{CH_2}$ SO_3H	3.02	¶
Compound *B*: CH_2OH CHOH $\mathit{CH_2}$ OH	3.57	¶

3.91 (compound *B*) resulted in the collapse to AB quartets of the two eight-line systems centred at δ 3.63 and δ 3.70. Irradiation at δ 3.63 or δ 3.70 caused simplification of the signal at δ 3.91. The remainder of the spectrum of compound *B*, which was almost identical to that of compound *A*, remained unaltered by these spin-decoupling experiments.

It followed that each compound could be regarded in two parts—a three-carbon moiety in which the differences between compounds *A* and *B* lay, and the remainder of the molecule (shown by ^{13}C NMR, see below, to contain seven carbon atoms). Both ^1H NMR and ^{13}C NMR (Table 2, compound *A* only) demonstrated the identity of the three-carbon moiety as 2,3-dihydroxypropane sulphonic acid in compound *A* and 1,2,3-trihydroxypropane in compound *B*. These identifications were confirmed by TLC (compound *A*) and HPLC (compound *B*) examination of the acid hydrolysates. The part of the 270-MHz ^1H NMR spectra which was identical for compounds *A* and *B* and assigned to the seven-carbon moiety contained signals attributable to two methyl groups and one methylene attached to arsenic. The rest of the spectrum (due to CH signals) was consistent with the presence of a sugar system. The 20.1-MHz ^{13}C NMR data for compound *A* confirmed the presence of a $C_{10}H_{17}$ skeleton (two CH_3 groups, three CH_2, five CH). Shieldings of the two methyl groups and the one methylene indicated that they were bound to arsenic. Shieldings of two methylenes and one methine were consistent with the presence of the sulphopropanediol residue as discussed above. The remainder of the methine shieldings demonstrated the presence of a furanoside with a transoid substitution arrangement at C-1 and C-2 (ref. 7). Of the possible candidates (β-xylo-, α-arabino-, β-ribo- and α-lyxo-), the closest match to the observed shieldings was obtained for a β-ribo- system when the effects on C-4 and C-3 of the postulated 5-deoxyfuranoside structure had been taken into account. However, the shieldings of an α-lyxoside were not sufficiently different from the observed values to eliminate completely an α-lyxo- in favour of a β-riboside system. Hydrolysis with aqueous sodium hydoxide resulted in the release of dimethylarsinic acid from both compounds *A* and *B*. Heating compound *A* with concentrated hydrochloric acid also resulted in the production of dimethylarsinic acid. In both experiments the remainder of the molecule was too far degraded for chemical identification of the sugar.

---

* The integrations of the signals were in accordance with their assignments.

† Only results of the most informative decoupling experiments are shown. In each of these no other part of the spectrum was altered. Irradiation, subsequently, at the chemical shifts of the simplified signals in these experiments resulted in the expected alterations to the spectrum at δ 4.25 (compound *A*) and δ 4.27 and δ 3.9 (compound *B*), but such changes were not amenable to simple analysis. Irradiation at the chemical shifts of other signals either had no effect or the altered signals were too close to the position of irradiation for reliable interpretaion.

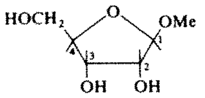
‡ The assignments of CH2 and CH3 are uncertain and may be reversed.

§ Not amenable to first order analysis.

‖ In furanoid systems where H-1 and H-2 are transoid the observed coupling constant approaches zero (ref. 15).

¶ Recorded at 60 MHz and not amenable to first order analysis.

Table 2 20.1 MHz ^{13}C NMR data for compound *A* and model compounds

Compound *A*			Model compounds		
Carbon*	Position	Shielding (p.p.m. relative to $(CH_3)_4Si$)	Structure with relevant carbons indicated (*italic*)		Shielding (p.p.m. relative to $(CH_3)_4Si$)
Two CH_3 groups	6, 7	14.7, 15.1	$(\mathit{C}H_3)_2AsO_2H$		17.1
CH	1	108.2		C-1	108.0§
CH	2	74.8		C-2	74.3§
CH	3	76.2†		C-3	70.9§
CH	4	77.3†		C-4	83.0§
CH_2	5	36.5	HO CH_2 $\mathit{C}H_2AsO_3H_2$		36.6
CH_2	1′	71.3‡	HO $\mathit{C}H_2$ CHOH CH_2SO_3H		65.0
CH	2′	66.9‡	HO CH_2 *C*HOH CH_2SO_3H		68.4
CH_2	3′	54.3	HO CH_2 CHOH $\mathit{C}H_2SO_3H$		54.0

* CH_3, CH_2 and CH groups were identified by off resonance decoupling.

† The differences in shieldings for C-3 and C-4 of compound *A* and the C-3 and C-4 of methyl β-ribofuranoside are accommodated by the 5-deoxy structure postulated for the former. It is assumed that the less electronegative substituent ($-CH_2As$) at the 5 position will produce effects on the shieldings of C-3 and C-4 analogous to (for example) the differences between the shieldings of C-4 and C-5 in glucose and 6-deoxyglucose[16]. Thus, differences in shielding of about 5 p.p.m. for C-3 and C-4 (in opposite directions) when compared with methyl β-ribofuranoside are accommodated.

‡ The differences in shieldings for C-1′ and C-2′ and the relevant carbons in the model compound are accounted for by the formation of the glycosidic linkage through $-C1'H_2O-$.

§ Ref. 7.

Compounds *A* and *B* are quite similar to isofloridoside (2,3-dihydroxypropyl α-D-galactopyranoside), a constituent of some red algae[8]. Free 2,3-dihydroxy-1-propane sulphonic acid has been reported as a component of algae[9]. However, compounds *A* and *B* are apparently unrelated to the arsenophospholipids and related de-acylated compounds containing a trimethylarsonium lactate residue demonstrated by Cooney *et al.*[10] to be present in unicellular marine algae exposed to radiolabelled arsenate. Such compounds were not detected in *Ecklonia.*

Evidently, *Ecklonia* is processing its arsenate uptake by conversion to the sugar derivatives described here. Benson has speculated[11] that algae have evolved the capacity to detoxify arsenate in marine waters low in phosphate. Western Australian coastal waters are notably deficient in phosphate[12]. The two compounds reported here can readily be envisaged as being intermediates in the production of dimethylarsinic acid (shown by Andreae[5] to be present in marine waters of the photic zone) and arsenobetaine. The latter would require further reduction and methylation of the arsenic atom and oxidation at the C-4 of the sugar residue. Such conversions might well be mediated by the microflora of the sediments contributed to by decaying *Ecklonia,* and arsenobetaine may enter the food chains, eventually reaching rock lobsters and whiting via detritovores. We are now studying these processes.

We thank Dr L. T. Byrne for ^{13}C NMR spectra, Dr A. J. Jones for 270-MHz ^{1}H NMR spectra and discussions, Drs D. A. Hancock and H. E. Jones for their interest, and Professor D. W. Cameron for much helpful discussion and encouragement. C, H microanalyses were performed by the Australian Microanalytical service.

Received 24 October; accepted 1 December 1980.

1. Penrose, W. R. *CRC crit. Rev. environ. Control* **4,** 465–479 (1974).
2. Edmonds, J. S. *et al. Tetrahedron Lett.* 1543–1546 (1977).
3. Cannon, J. R., Edmonds, J. S., Francesconi, K. A. & Langsford, J. B. *Management and Control of Heavy Metals in the Environment, Int. Conf. Lond.*, 283–286 (CEP Consultants, Edinburgh, 1979).
4. Edmonds, J. S. & Francesconi, K. A. *Mar. Pollut. Bull.* (in the press).
5. Andreae, M. O. *Deep Sea Res.* **25,** 391–402 (1978).
6. Merijanian, A. & Zingaro, R. A. *Inorg. Chem.* **5,** 187–191 (1966).
7. Ritchie, R. G. S., Cyr, N., Korsch, B., Koch, H. J. & Perlin, A. S. *Can. J. Chem.* **53,** 1424–1433 (1975).
8. Lindberg, B. *Acta chem. scand.* **9,** 1097–1099 (1955).
9. Busby, W. F. *Biochim. biophys. Acta* **121,** 160–161 (1966).
10. Cooney, R. V., Mumma, R. O. & Benson, A. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4262–4264 (1978).
11. Maugh, II, T. H. *Science* **203,** 637 (1979).
12. Rochford, D. J. *Aust. J. mar. Freshw. Res.* **13,** 226–251 (1962).
13. Massey, A. G., Randall, E. W. & Shaw, D. *Spectrochim. Acta* **A23,** 1235–1242 (1967).
14. Lemieux, R. U. & Stevens, J. D. *Can. J. Chem.* **44,** 249–262 (1966).
15. Inch, T. D. *A. Rev. NMR Spectrosc.* **2,** 35–82 (1969).
16. Gorin, P. A. J. & Mazurek, M. *Can. J. Chem.* **53,** 1212–1223 (1975).

Active site and catalytic mechanism of phospholipase A_2

Bauke W. Dijkstra, Jan Drenth & Kor H. Kalk

Laboratory of Chemical Physics, University of Groningen,
Nijenborgh 16, 9747 AG Groningen, The Netherlands

The esterolytic enzyme phospholipase A_2 specifically splits the 2-acyl linkage of phosphoglycerides[1] in a calcium-dependent reaction. In the pancreas the enzyme occurs as a zymogen which is activated on secretion into the duodenal tract by the removal of seven amino acid residues from the N terminus by trypsin[2,3]. Having refined our X-ray analysis of the crystal structure of bovine pancreatic phospholipase A_2 from 2.4 Å (ref. 4) to 1.7 Å resolution[5], we now describe how the structure of the molecule may account for the specificity of the enzyme and for the sudden and dramatic change in activity when the substrate concentration passes the critical micelle concentration[6].

Analysis of the structure at 2.4 Å resolution[4] showed that the active site is in a depression at the molecular surface. The essential residue, His 48, which is directly involved in catalysis[7], is located in this region (Fig. 1). The ε-N atom of His 48 is in close contact with one of the carboxylate oxygens of Asp 99 (for residue numbering see ref. 8). The wall of the active site is covered with hydrophobic residues: Phe 5, Ile 9, Phe 22, Ala 102 and 103, Phe 106 and the disulphide bridge between Cys 29 and Cys 45. All these residues are invariant in the phospholipase A_2 sequences known so far, except for an occasional replacement of Ala 103 by a Val (ref. 9) and Phe 22 by a Tyr (refs 9–13). The essential calcium ion is also in this region, but near Asp 49, not Asp 99 as previously suggested[4].

The interior of the active site mainly consists of invariant hydrophobic residues, whereas the entrance is composed of variable residues, hydrophobic and polar (Fig. 2).

Although Asp–His couples occur frequently in protein molecules, the presence of this couple in the active site of phospholipase A_2 suggests a similarity to the proton relay system in the proteolytic serine enzymes. In these proteases—as in all proteolytic enzymes[14,15]—there are three main stages in the splitting of the peptide bond: (1) nucleophilic attack by the serine OH; (2) proton transfer by the histidine's imidazole ring; and (3) fixation and stabilization of the substrate's carbonyl oxygen by slightly positively charged NH groups. In deacylation a water molecule replaces the serine OH group as the nucleophile. Phospholipase A_2 has exactly the same configuration in its

0028-0836/81/070604-03$01.00

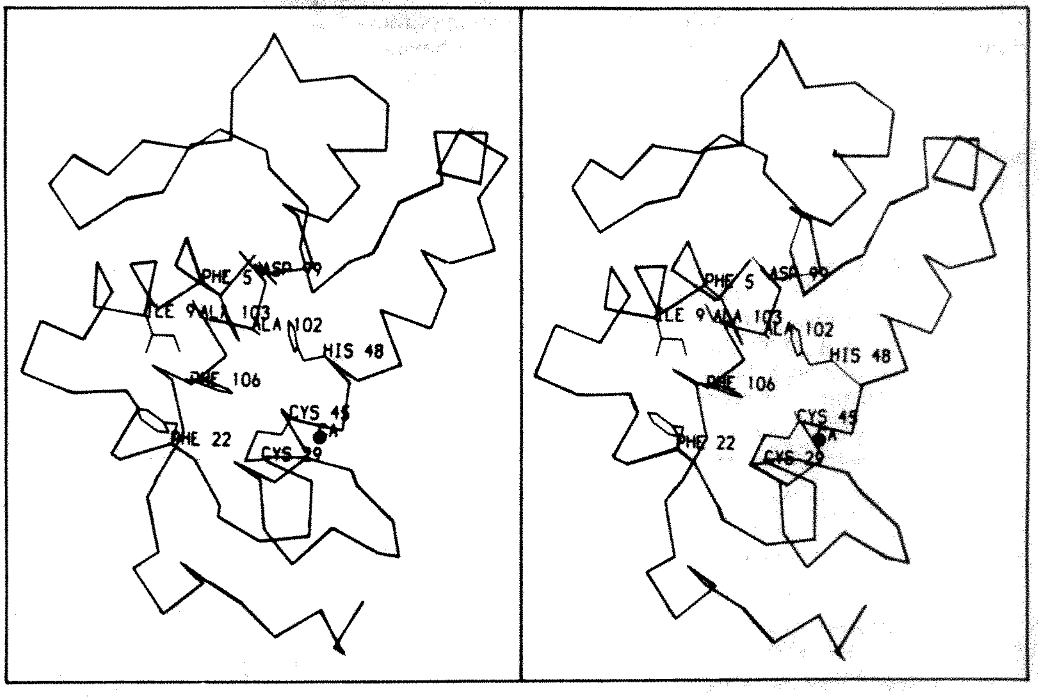


Fig. 1 A stereo pair showing the essential active site residues. The black dot is the calcium ion.

active site but has a water molecule instead of a serine OH. This water molecule is in the plane of the imidazole ring 3.1 Å from His 48δ-N1. Thus, the ester hydrolysis by phospholipase A_2 seems to be identical to the deacylation step—which is also ester hydrolysis—in the serine proteases, the only difference being that in the latter the substrate is covalently bound to the enzyme. Fixation and stabilization of the carbonyl oxygen atom of the ester bond might occur in phospholipase by means of the NH of Gly 30 and possibly by the calcium ion. This ion is also available for binding the substrate's phosphate group, especially because it has only one carboxyl group (Asp 49) in its coordination sphere. Substrate binding studies are required to test this hypothesis experimentally.

Phospholipase A_2 has a much higher activity towards aggregated than single phospholipid molecules (up to 1,000, ref. 16). Some workers attribute this effect to an entropic factor—a preferred orientation of the enzyme with its active site directed to the substrate layer[17], or a preferred substrate conformation[16]. Others explain the effect by a change in enzyme conformation at the active site as a result of the binding to the substrate layer[6,18]. With only the enzyme's native structure available the conformational change cannot be verified. However, we do know the part of the enzyme's molecular surface which interacts with the micelle and also that a complementarity exists between the active site and the conformation of the substrate molecules in the aggregated state.

If the enzyme has its active site directed towards the micellar surface three distinct regions for interaction with the micelle can be discerned (Fig. 2). One region comprises residues in a ring surrounding the entrance to the active site, the other two regions are at the extremes of the molecule. Several residues in these regions—those at positions 3, 6, 19, 31 and 69 (ref. 19)—have been implicated in the interaction with micellar substrates. Although none of these regions shows residues conserved at homologous positions, the hydrophobic or polar character of these residues is maintained. A special feature is the excess of at least two basic over acid residues, suggesting salt bridge formation with the phosphate groups in the micelle which would thus orient the enzyme with respect to the micelle surface. This is in agreement with a proposal by Brockerhoff and Jensen[17].

Crystal structure and NMR studies have shown that phospholipid molecules in micelles or pseudo-micelles have a unique conformation[20]. Their main features are parallel and extended fatty acid chains, with a kink in the second chain at the α-C atom. The shape of the phospholipase A_2 active site is such that a substrate molecule in this conformation can easily enter the active site and approach His 48 with its scissile bond. A conformational change in the substrate molecule is not required, and model building studies indicate that the kink is an absolute requirement for fitting. The calcium ion in that case is available for binding the phosphate group, and the active site holds between four and six CH_2 groups of each fatty acid chain. This shows—as previously suggested[16]—that besides the orientation of the enzyme molecule a second entropic advantage is the presence of the right substrate conformation in aggregates like micelles.

Our crystallographic studies have also analysed the structure of bovine pancreatic prophospholipase A_2. This precursor does not show the rate increase towards aggregated substrate. Although the structure of porcine pancreatic prophospholipase A_2 at 3 Å resolution has been reported[15], it was later

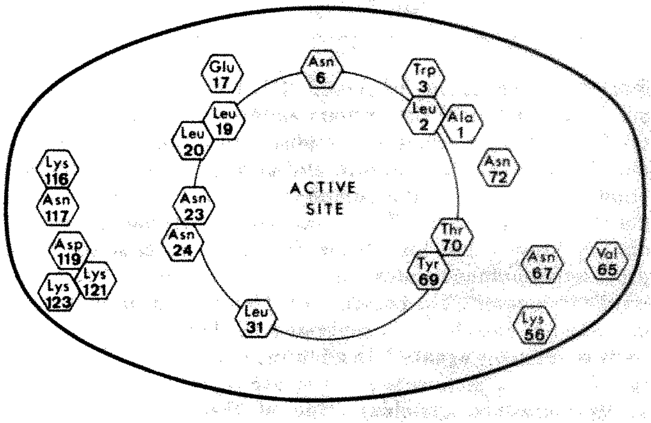


Fig. 2 Schematic representation of the face of the molecule which is directed towards the phospholipid layer. The residues in boxes, in the bovine enzyme, have their side chain pointing to the substrate layer. Residue 1 is incorporated because of its free NH_2 which is essential for binding. One set of residues is in a ring surrounding the entrance to the active site. Two other sets are on the left and on the right.

found to have been based on an incorrect interpretation of the electron density map. Recently, high-quality crystals of bovine prophospholipase A_2 have become available, and preliminary crystallographic results indicate differences with the enzyme itself at the N terminus due to an increased mobility in this region. This agrees with studies by Slotboom *et al.*[21] showing that a precise architecture of this region is required for micelle binding.

We thank Professor G. H. de Haas and his coworkers for their cooperation, Dr W. G. J. Hol for stimulating discussions and Mrs C. A. Torfs-van Cotthem for her contributions. This work was supported in part by the Netherlands Foundation for Chemical Research with financial aid from the Netherlands Organisation for the Advancement of Pure Research.

Received 21 July; accepted 5 December 1980.

1. Van Deenen, L. J. M. & De Haas, G. H. *Adv. Lipid Res.* **2,** 167–234 (1964).
2. De Haas, G. H., Postema, N. M., Nieuwenhuizen, W. & Van Deenen, L. L. M. *Biochim. biophys. Acta* **159,** 103–117 (1968).
3. Dutilh, C. E., Van Doren, P. J., Verheul, F. E. A. M. & De Haas, G. H. *Eur. J. Biochem.* **53,** 91–97 (1975).
4. Dijkstra, B. W., Drenth, J., Kalk, K. H. & Vandermaelen, P. J. *J. molec. Biol.* **124,** 53–60 (1978).
5. Dijkstra, B. W., Drenth, J. & Kalk, K. H. *J. molec. Biol.* (submitted).
6. Pieterson, W. A., Vidal, J. C., Volwerk, J. J. & De Haas, G. H. *Biochemistry* **13,** 1455–1460 (1974).
7. Verheij, H. M. *et al. Biochemistry* **19,** 743–750 (1980).
8. Meyer, H., Verhoef, H., Hendriks, F. F. A., Slotboom, A. J. & De Haas, G. H. *Biochemistry* **18,** 3582–3588 (1979).
9. Fohlman, J., Lind, P. & Eaker, D. *FEBS Lett.* **84,** 367–371 (1977).
10. Heinrikson, R. L., Krueger, E. T. & Keim, P. S. *J. biol. Chem.* **252,** 4913–4921 (1977).
11. Halpert, J. & Eaker, D. *J. biol. Chem.* **250,** 6990–6997 (1975); **251,** 7343–7347 (1976).
12. Kondo, D., Narita, K. & Lee, C. Y. *J. Biochem., Tokyo* **83,** 101–115 (1978).
13. Botes, D. P. & Viljoen, C. C. *J. biol. Chem.* **249,** 3827–3835 (1974).
14. Drenth, J. *Recl Trav. chim. Pays-Bas* **99,** 185–190 (1980).
15. Drenth, J., Enzing, C. M., Kalk, K. H. & Vessies, J. C. A. *Nature* **264,** 373–377 (1976).
16. Wells, M. A. *Biochemistry* **13,** 2248–2257 (1974).
17. Brockerhoff, H. & Jensen, R. G. *Lipolytic Enzymes,* 194–243 (Academic, London, 1974).
18. Verger, R. & De Haas, G. H. *A. Rev. Biophys. Bioengng* **5,** 77–117 (1976).
19. Meyer, H. *et al. Biochemistry* **18,** 3589–3597 (1979).
20. Hauser, H., Guyer, W., Pascher, I., Skrabal, P. & Sundell, S. *Biochemistry* **19,** 366–373 (1980).
21. Slotboom, A. J., Van Dam-Mieras, M. C. E. & De Haas, G. H. *J. biol. Chem.* **252,** 2948–2951 (1977).

Assignment of the disulphide bonds of leukocyte interferon

Ronald Wetzel

Genentech Inc., 460 Point San Bruno Boulevard,
South San Francisco, California 49080

Nucleotide sequencing of cloned cDNA can lead to rapid and precise prediction of the primary amino acid sequence of gene products, but it cannot establish such post-translational modifications as the existence and arrangement of disulphide bonds. For example, the complete sequences of at least one fibroblast[1] and seven[2–4,12] leukocyte interferon genes are already known, but knowledge derived from analysis of the proteins is confined to information on the N-termini[5–8] and some tryptic fragments[9]. The presence of at least one disulphide bond in leukocyte interferon is suggested by that molecule's sensitivity to reducing agents[10]. In addition, comparison of all leukocyte interferon gene sequences so far reported indicates four highly conserved cysteines. One of these genes has been engineered for efficient direct expression in *Escherichia coli*[4] and we have purified the gene product, leukocyte interferon A (IFN-αA, Fig. 1) from bacterial extracts to a single species of molecular weight (MW) 19,400[13]. I report here the determination of the disulphide bonds of the purified protein by analysis of tryptic fragments. The results indicate that Cys 1 is bonded to Cys 98, and Cys 29 is bonded to Cys 138.

```
                                    10                                      20
CYS ASP LEU PRO GLN THR HIS SER LEU GLY SER ARG ARG THR LEU MET LEU LEU ALA GLN MET ARG LYS ILE SER

                30                                      40                                      50
LEU PHE SER CYS LEU LYS ASP ARG HIS ASP PHE GLY PHE PRO GLN GLU GLU PHE GLY ASN GLN PHE GLN LYS ALA

                                    60                                      70
GLU THR ILE PRO VAL LEU HIS GLU MET ILE GLN GLN ILE PHE ASN LEU PHE SER THR LYS ASP SER SER ALA ALA

                80                                      90                                      100
TRP ASP GLU THR LEU LEU ASP LYS PHE TYR THR GLU LEU TYR GLN GLN LEU ASN ASP LEU GLU ALA CYS VAL ILE

                                    110                                     120
GLN GLY VAL GLY VAL THR GLU THR PRO LEU MET LYS GLU ASP SER ILE LEU ALA VAL ARG LYS TYR PHE GLN ARG

                130                                     140                                     150
ILE THR LEU TYR LEU LYS GLU LYS LYS TYR SER PRO CYS ALA TRP GLU VAL VAL ARG ALA GLU ILE MET ARG SER

                                    160
PHE SER LEU SER THR ASN LEU GLN GLU SER LEU ARG SER LYS GLU
```

Fig. 1 The amino acid sequence of IFN-αA predicted from the nucleotide sequence of the gene in the plasmid directing its expression in *E. coli*[4]. The cysteine residues are underlined.

Digestion of native IFN-αA with trypsin produces a mixture of fragments which can be partially resolved by HPLC as shown in Fig. 2*a*. Treatment of part of the tryptic digest before HPLC analysis with dithiothreitol (DTT) yields the profile shown in Fig. 2*b*. At least five significant peaks are lost on DTT reduction, indicating the presence of a disulphide bond in each. Further experiments showed that peak 1 is associated with one of the disulphide bonds and peaks 2–5 are associated with the other.

At increasing trypsin/interferon ratios, peaks 2 and 3 increase in size relative to peaks 4 and 5 indicating that they are related by the presence of a resistant trypsin cleavage site. At a trypsin/interferon ratio of 1:10, peaks 4 and 5 almost disappear. The fact that the 2, 3 doublet and the 4, 5 doublet wax or wane in unison suggests that 2 is related to 3 and 4 is related to 5.

Samples of peaks 1, 4 and 5 (3–12 μg) were subjected to amino acid analysis. The results were compared with calculated values for all possible disulphide-containing peptides predicted from the nucleotide sequence of the gene (Table 1). Also

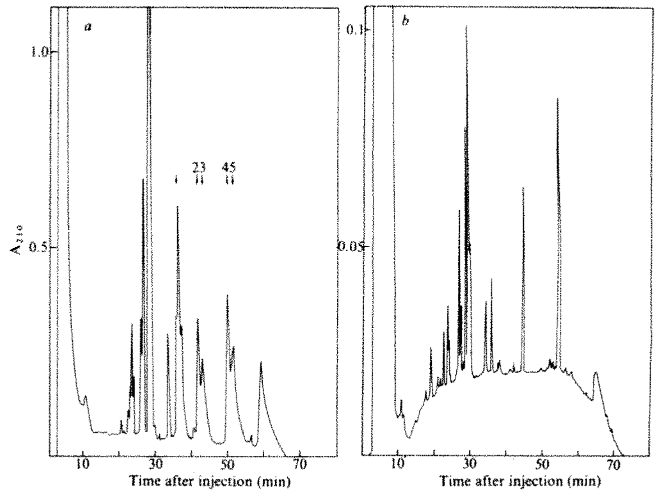


Fig. 2 HPLC of a trypsin digest of IFN-αA. A solution of 500 μg of native IFN-αA in 400 μl of 50 mM sodium phosphate, *p*H 7.8 was incubated with 8 μg of TPCK-trypsin (Worthington) for 24 h at room temperature. Analysis was on an Altex ODS-Ultrasphere column with a linear gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid[11]. *a*, One half of the digest was injected and individual peaks collected. *b*, 20 μl of the digest were added to 20 μl of 7 M urea, 50 mM DTT and incubated for 2 h at room temperature before injection on the same HPLC system.

0028-0836/81/070606-02$01.00

Table 1 Mole percent amino acid compositions of cysteine-containing tryptic peptides

Amino acid	Calculated values for possible peptides									Values obtained		
	1–29	1–98	1–138	29–98	29–138	98–138	1–98⁺	29–98⁺	98⁺–138	Peptide 1	Peptide 4	Peptide 5
Ala	0	2.56	5.26	2.86	6.67	5.56	5.88	6.38	8.33	5.83	5.63	5.80
Arg	5.56	2.56	10.53	0	6.67	2.78	1.96	0	2.08	6.51	1.87	2.20
Asx	5.56	7.69	5.26	5.71	0	5.56	11.76	10.64	10.42	0	12.67	12.00
Cys/2	(2)*	(2)*	(2)*	(2)*	(2)*	(2)*	(2)*	(2)*	(2)*	n.d.	n.d.	n.d.
Glx	5.56	17.95	10.53	17.14	6.67	19.44	15.69	14.89	16.67	13.37	16.31	16.87
Gly	5.56	7.69	5.26	5.71	0	5.56	5.88	4.26	4.17	1.88	6.38	6.77
His	5.56	2.56	5.26	0	0	0	1.96	0	0	0	1.63	1.94
Ile	5.56	2.56	0	5.71	6.67	2.78	1.96	4.26	2.08	6.51	2.30	1.70
Leu	22.22	15.38	10.53	17.14	13.33	11.11	15.69	17.02	12.50	13.24	15.03	14.35
Lys	5.56	2.56	0	5.71	6.67	2.78	3.92	6.38	4.17	6.60	4.11	4.15
Met	0	2.56	0	2.86	0	2.78	1.96	2.13	2.08	0	1.54	1.42
Phe	5.56	2.56	0	5.71	6.67	2.78	1.96	4.26	2.08	5.86	1.98	2.30
Pro	5.56	5.13	10.53	2.86	6.67	5.56	3.92	2.13	4.17	8.34	4.84	4.18
Ser	22.22	5.13	15.79	5.71	20.00	2.78	7.84	8.51	6.25	15.94	6.41	8.15
Thr	5.56	10.26	5.26	8.57	0	8.33	9.80	8.51	8.33	0	9.56	9.39
Trp	0	0	(1)*	0	(1)*	(1)*	(1)*	(1)*	(2)*	n.d.	n.d.	n.d.
Tyr	0	5.13	5.26	5.71	6.67	8.33	3.92	4.26	6.25	5.61	4.36	3.23
Val	0	7.69	10.53	8.57	13.33	13.89	5.88	6.38	10.42	10.27	5.36	5.50

Cystine peptides are named according to the cysteine residues they contain, with 98^+ denoting the Cys 98 peptide obtained when the Lys–Phe bond after Lys 83 is not cleaved. Analyses were obtained from AAA Labs, Mercer Island, Washington.

* Values for these amino acids were not used in calculating compositions because cysteine and tryptophan were not determined in the amino acid analysis.

included as possible structures were the peptides which would be generated from the lack of a trypsin cut at the Lys–Phe bond after Lys 83. Analysis of the results indicates that peak 1 corresponds to a peptide with a disulphide bond between Cys 29 and Cys 138, and peaks 4 and 5 correspond to a peptide with a disulphide bond between Cys 1 and Cys 98 which was not cleaved by trypsin after Lys 83. Comparison of the analysis of peak 5 with that of peak 4 shows several minor differences, which may be attributed to experimental error due to small sample size. Peaks 4 and 5 could differ in the extent to which they have undergone deamidation of amide side chains. Such heterogeneity would not be unexpected, and would not show up in the analysis of an acid hydrolysate.

The presence of the two disulphide bonds, one of them involving Cys 1, suggests that an early cysteine should be present in the mature sequences of other leukocyte interferons, a postulate contradicted by some published N-terminal protein sequence data[5,8] but consistent with all reported DNA sequences. Since two of the three cysteines of fibroblast interferon[1] are conserved in the sequence of leukocyte interferon, and since these two conserved cysteines form one of the disulphides in leukocyte interferon, it seems likely that a disulphide bond exists between Cys 31 and Cys 141 of the fibroblast molecule. Regardless of the presence and arrangement of a disulphide bond in fibroblast interferon, there can be no more than one, and this work thus demonstrates a structural difference between the two kinds of acid-stable interferons.

Received 4 November, accepted 8 December 1980

1 Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. Gene 10, 11–15 (1980)
2 Mantei, N. et al. Gene 10, 1–10 (1980).
3. Streuli, M., Nagata, S. & Weissmann, C. Science 209, 1343–1347 (1980).
4 Goeddel, D. V. et al. Nature 287, 411–416 (1980)
5 Zoon, K. C. et al. Science 207, 527–528 (1980).
6 Taira, H. et al. Science 207, 528–530 (1980)
7 Knight, E. Jr, Hunkapiller, M. W., Korant, B. D., Hardy, R. W. F. & Hood, L. E. Science 207, 525–526 (1980)
8 Levy, W. P., Shively, J., Rubinstein, M., Del Valle, U. & Pestka, S. Proc. natn. Acad. Sci. U.S.A. 77, 5102–5104 (1980)
9 Allen, G. & Fantes, K. H. Nature 287, 408–411 (1980)
10 Mogensen, K. E. & Cantell, K. J. gen. Virol. 22, 95–103 (1974).
11 Bennett, H. P. J., Browne, C. A., Goltzman, D. & Solomon, S. in Proc. 6th American Peptide Symp. (eds Gross, E. & Meienhofer, J.) 121–124 (Pierce, Rockford, 1979).
12. Goeddel, D. V. et al. Nature (in the press).
13. Wetzel, R. et al. J. Interferon Res. (in the press).

Unique transforming gene in carcinogen-transformed mouse cells

Ben-Zion Shilo & Robert A. Weinberg

Center for Cancer Research and Department of Biology,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139

Observations from various biological systems suggest that chemical carcinogens transform cells by causing somatic mutations through interacting with DNA[1–5]. These observations raise several fundamental questions regarding the nature of the transforming genes that are the target of the carcinogens. For example, are there unique target gene(s) for chemical carcinogens, or are there multiple potential targets, any of which, when altered, can induce the transformed phenotype? This problem has been approached previously by comparing the rate of transformation by chemical carcinogens with the rate of appearance of dominant mutations in unique cellular genes encoding selectable markers. The results suggested a limited number of targets for transformation, since transformation occurred at a rate that was only 10–100 times greater than the appearance of ouabain-resistant mutants[6–8]. These results, however, are only rough estimates, because the optimal conditions for induction of transformation and the induction of mutations are known to differ[6]. In addition, lesions other than point mutations caused by carcinogens, such as frameshifts, may induce transformation efficiency but not ouabain resistance. We now present an alternative experimental approach to estimating the number of targets for chemical carcinogens, based on the fact that DNAs of certain chemically transformed cell lines can induce the appearance of foci when transferred on to NIH3T3 cells[9]. These DNAs were treated with a variety of restriction endonucleases before testing their transforming activities by DNA transfer. We have found the same pattern of sensitivity and resistance to inactivation by different endonucleases in all DNAs studied, suggesting that the same transforming gene is being transferred in all cases.

0028–0836/81/070607–03$01.00

Table 1 The effect of restriction enzyme cleavages on the transfections of DNAs of 3-MC-transformed cells

DNA source	Origin of donor line	No enzyme	Foci				
			*Bam*HI	*Eco*RI	*Hin*dIII	*Xho*I	*Sal*I
MCA16-5	A focus derived from transfection of DNA from the 3-MC-transformed C3H10T1/2 line MCA16 (ref. 17)	73, 58	14, 25	0	0	33	50
MCA5	3-MC-transformed C3H10T1/2 cells[18]	5, 6, 8	7	0	0	4	8
MB66MCAad36	3-MC-transformed C3H10T1/2 cells	15	14, 9, 9	0	0	34	6
MC5-5-6	A focus derived from transfection of chromosomes from the 3-MC-transformed BALB3T3 line MC5-5 (ref. 19)	60	17	0	0	6	5
NFS-1	A focus from transfection of NFS mouse liver DNA	17, 26	0	0	ND	ND	ND
NIH3T3	NIH3T3 cells	0	0	ND	ND	0	ND

In each transfection 75 μg of DNA were applied to 1.5×10^6 NIH3T3 cells, and the appearance of foci was scored 14 d later. The transfection efficiency of the DNA following digest with enzymes that are non-cutters is occasionally reduced by a factor of two to three. We do not understand the basis for this reduction. ND, not determined.

DNA-mediated gene transfer has recently been applied to the study of chemical carcinogenesis[9]. DNAs from a variety of chemically transformed mouse lines were applied to NIH3T3 cells, and their ability to induce foci of transformed cells was scored. DNAs extracted from five lines that were independently transformed *in vitro* by 3-methylcholanthrene (3-MC) were shown to induce the appearance of foci[9], thus demonstrating directly that the DNA of these chemically transformed cells was altered relative to the DNAs of untransformed mouse cells. These alterations were probably of central importance in carcinogenesis, as they resulted in the generation of potent transforming genes not detectable in normal cellular DNA.

The naked DNA being introduced into NIH3T3 cells in these transfection assays can be biochemically modified before its transfer by treatment with site-specific endonucleases. An enzyme that cuts within a functionally essential region of the transforming gene would be expected to abolish the transforming ability of the DNA. Conversely, enzymes cleaving only outside the functionally important regions of a gene should leave the biological activity of the DNA unaffected. It is also essential for the initiation of molecular cloning of the gene[10] to obtain information about enzymes that are 'cutters' or 'non-cutters' of the transforming activity.

Because of the random occurrence of cleavage sites for restriction enzymes, one cannot predict the presence of such sites within a given gene. This randomness confers on each gene a specific pattern of sensitivity or resistance to inactivation by a series of restriction enzymes applied to the DNA encoding the gene. It is highly unlikely that two genes of unrelated sequence will have an identical spectrum of sensitivities and resistances to many restriction enzymes. Such enzyme cleavage before transfection, therefore, provides a useful method for comparing the structures of transforming genes being transferred from the different chemically transformed donor cell lines. We have used this procedure to analyse transforming genes of four of the 3-MC-transformed mouse cell lines, either by cleavage of the DNAs of the transformed donor cells or by cleavage of the DNAs of foci derived from transfection of the donor cell DNAs. This latter step was made possible by our earlier demonstration that the transforming genes can be serially passaged from donors to recipients through several cycles of transfection[9].

Before transfection, each of the four DNAs was cleaved with one of a series of five restriction enzymes. Each of these enzymes recognizes a different hexanucleotide sequence at its site of cleavage. In every case the completeness of cleavage by an enzyme was determined by taking an aliquot of the reaction mixture, adding a small amount of a DNA of defined sequence and monitoring the digestion products of that sequence by gel electrophoresis. An example of this assay is shown in Fig. 1. In every transfection, a positive control of uncleaved DNA was included in parallel. We demanded significant numbers of foci from cultures transfected with the positive control DNA before regarding the results of a transfection as interpretable.

The results of these experiments are summarized in Table 1. The data indicate an identical pattern of sensitivity and resistance to the five restriction enzymes among the DNAs of the four lines tested. We conclude that the transforming genes of the four mouse lines have a high probability of being associated with an identical set of nucleotide sequences. The interpretation of these results is even more compelling if the average cleavage frequencies in mammalian DNA of the different enzymes used are taken into account. *Bam*HI, *Eco*RI and *Hin*dIII cleave mammalian DNA into fragments which average 4 kilobase pairs in mass. Among the enzymes of a similar cutting frequency tested (including *Bgl*II and *Xba*I for MCA16-5 DNA; data not shown), only *Bam*HI did not inactivate the transforming activity of the DNA. It was thus striking that the same frequently cutting enzyme was not a cutter for any of the four DNAs.

The following statistical considerations support the conclusion that the same transforming gene is being transferred in the four cases. The probability that a gene in the 10-kilobase pair size range would be inactivated by cleavage with an enzyme that cuts on average every 4 kilobase pairs is about 0.8. Conversely, the probability of that gene being inactivated by an enzyme that cuts every 30 kilobase pairs (such as *Xho*I or *Sal*I) is only 0.3. Thus, the probability of two lines displaying the same pattern we observed for the five enzymes would be $0.2(0.8)^2(0.7)^2 = 0.06$. The probability of four lines displaying the same pattern would be $(0.06)^3 = 2 \times 10^{-4}$. Therefore, the chances of the results we observed occurring coincidentally are as low as 2×10^{-4}.

The results derived from the non-cutting enzymes such as *Bam*HI might be interpreted differently—that these non-cutting enzymes did destroy the activity of the gene observed on transfection of an uncleaved DNA, while at the same time creating a novel, equally potent transforming gene as a direct consequence of cleavage of normal cellular sequences at a critical target site. This possibility becomes more likely in the light of recent work suggesting that random breakage of normal cellular DNA induced by sonication was able to create at a low frequency novel, transfectable transforming genes[11]. We therefore examined the effects of cleavage of normal mouse DNA (extracted from NIH3T3 cells) by *Bam*HI or *Xho*I. Because no foci were observed following transfection (Table 1), we concluded that cleavage by these enzymes does not activate latent transforming genes which may exist in normal cellular DNA.

Additional data support the conclusion that the spectrum of resistance and sensitivity to specific endonucleases is unique to the 3-MC-transformed lines. A transformed NIH3T3 cell line, which was derived after exposure to normal mouse liver DNA, possesses a transforming gene which is inactivated by *Bam*HI (Table 1), in contrast to the 3-MC-transformed lines. A rat neuroblastoma gene exhibits a totally different pattern of resistance and sensitivity to the enzymes tested here[12]. Finally, in a series of murine retrovirus genomes whose transforming genes have been characterized by transfection into NIH3T3 cells, each exhibits a unique pattern of inactivation by the different site-specific endonucleases[13–16]. The characteristic pattern of the 3-MC-transformed lines observed here is thus a reflection of the physical structure of the gene being transferred and is not due to an effect of these enzymes on the general process of transfection.

The present conclusions on the identity of the transforming genes in the four 3-MC-transformed lines will be strengthened when the respective transforming alleles are isolated as molecular clones. However, neither the present results nor the anticipated molecular cloning allow us to conclude that the transforming alleles observed in these cell lines were created as a direct result of exposure to 3-MC, the original carcinogen used. Rather, it is possible that in cells mutagenized by carcinogen, a second, independently occurring mutational event may occur on subsequent cell passage, which in turn results in the generation of the transforming genes described here.

Our conclusions regarding the multiplicity of potential transforming genes which can be activated by exposure to chemical carcinogens are limited to the spectrum of transformed lines that has been tested. They apply only to transformed mouse cells which are derived from exposure of monolayer fibroblast cultures to 3-MC *in vitro*. Because several 3-MC-transformed mouse fibroblast lines failed to yield biologically active DNA in transfections[9], the present results may be relevant only to a subset of 3-MC-transformed mouse fibroblasts. Nevertheless, the data suggest that the number of transforming genes induced by chemical carcinogens may be small, and that the molecular mechanisms of chemical transformation depend on activation of limited and unique pathways for transformation.

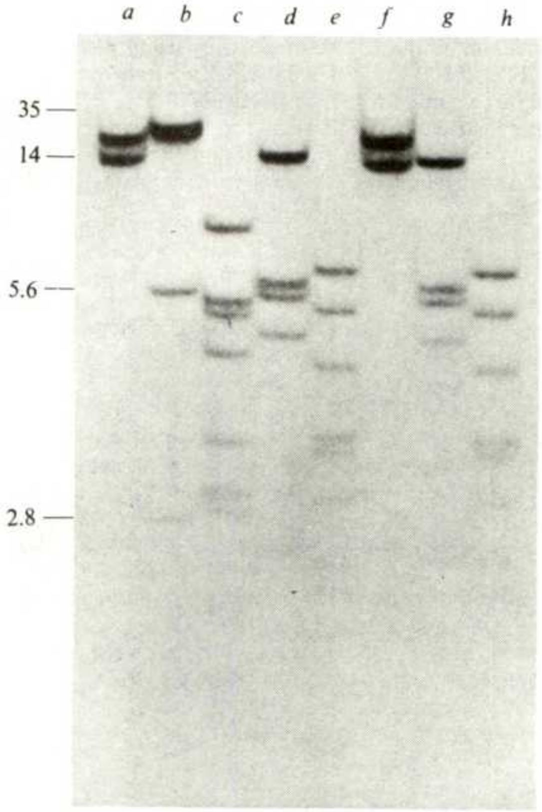


Fig. 1 Analysis of completeness of digestion of cellular DNAs by restriction endonucleases. Cellular DNA was incubated in restriction endonuclease buffer at a concentration of 20–50 μg ml^{-1}. After addition of the enzyme, a 50-μl aliquot was incubated with 50 pg of *in vivo* ^{32}P-labelled adenovirus 5 DNA (35 kilobase pairs). The reaction was incubated at 37 °C for 5 h and the aliquots run on a 1% agarose gel. The gel was fixed for 1 h in methanol, dried and exposed to X-ray film. The cleaved cellular DNA was treated with phenol and chloroform and ethanol precipitated before transjection. *a–e*, MB66MCA ad36 DNA digested with *Bam*HI, *Eco*RI, *Hin*dIII, *Xho*I and *Sma*I, respectively. *f–h*, MCA16-5 DNA digested with *Bam*HI, *Xho*I and *Sma*I, respectively. The molecular weights are given in kilobase pairs. The results of the transfection of the DNAs whose digests were monitored on this gel are given in Table 1.

We thank Chiaho Shih for his help in the transfections and for characterizing the rat neuroblastoma gene. This work was supported by the Rita Allen Foundation of which R.A.W. was a fellow and by NIH core grant CA14051 to S. Luria. B.S. is a Chaim Weizmann postdoctoral fellow.

Received 21 August; accepted 24 November 1980.

1. McCann, J., Choi, E., Yamasaki, E. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **72,** 5135–5139 (1975).
2. McCann, J. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **73,** 950–954 (1976).
3. Bridges, B. A. *Nature* **261,** 195–200 (1976).
4. Bouck, N. & diMayorca, G. *Nature* **264,** 722–727 (1976).
5. Barrett, J. C., Tsutsui, T. & Ts'o, P. O. P. *Nature* **274,** 229–232 (1978).
6. Landolph, J. R. & Heidelberger, C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 930–934 (1979).
7. Parodi, S. & Brambilla, G. *Mutat. Res.* **47,** 53–74 (1977).
8. Huberman, E., Mager, R. & Sachs, L. *Nature* **264,** 360–361 (1976).
9. Shih, C., Shilo, B., Goldfarb, M. P., Dannenberg, A. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5714–5718 (1979).
10. Shilo, B. *et al. 12th Miami Winter Symp.* (in the press).
11. Cooper, G. M., Okenquist, S. & Silverman, L. *Nature* **284,** 418–421 (1980).
12. Shih, C., Padhy, L. C., Murray, M. & Weinberg, R. A. (in preparation).
13. Andersson, P., Goldfarb, M. P. & Weinberg, R. A. *Cell* **16,** 63–75 (1979).
14. Canaani, E., Robbins, K. C. & Aaronson, S. A. *Nature* **282,** 378–383 (1979).
15. Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J. V. & Vande Woude, G. F. *Science* **207,** 1222–1224 (1980).
16. Ellis, R. W. *et al. J. Virol.* **35,** 79–92 (1980).
17. Reznikoff, C. A., Bertram, J. S., Brankow, D. W. & Heidelberger, C. *Cancer Res.* **33,** 3239–3249 (1973).
18. Rapp, U. R., Nowinski, R. C., Reznikoff, C. A. & Heidelberger, C. *Virology* **65,** 392–409 (1975).
19. Benveniste, R. E., Todaro, G. J., Scolnick, E. M. & Parks, W. P. *J. Virol.* **12,** 711–720 (1973).

Distribution of the insertion sequence IS1 in Gram-negative bacteria

Kate Nyman, Kenzo Nakamura*, Hisako Ohtsubo & Eiichi Ohtsubo

Department of Microbiology and *Biochemistry, Basic Health Sciences Center, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794

Translocation of DNA segments is a recombinational event seen in both eukaryotic and prokaryotic chromosomes, and it is thought to be involved in controlling gene expression and in the evolution of chromosomes[1,2]. In bacteria, insertion (IS) and transposable (Tn) elements[3] not only translocate their own DNA, but also promote the rearrangement of both bacterial chromosomes and the plasmic genomes carrying them. The insertion element IS1 is one such element which is 768 base pairs long[4–7]. IS1 is involved in the generation of deletion mutations and in the fusion of two different plasmid genomes[8–12]. It can also promote the translocation of DNA segments flanked by two copies of IS1 to give rise to transposable elements responsible for antibiotic resistance and enterotoxin production[13–16]. We report here the distribution of the IS1 sequence in various bacterial DNAs, particularly in the family Enterobacteriaceae. Comparison of the results with the phylogenetic relationship of these bacteria suggests that IS1 was transferred from one bacterium to another after their divergence and in some bacteria the copy number of IS1 increased by translocation. The increase in the number of copies of IS1 in bacteria may increase the probability of the genetic rearrangement responsible for the generation of resistance and enterotoxin plasmids, the existence of which is a serious problem in medical microbiology.

0028-0836/81/070609-04$01.00

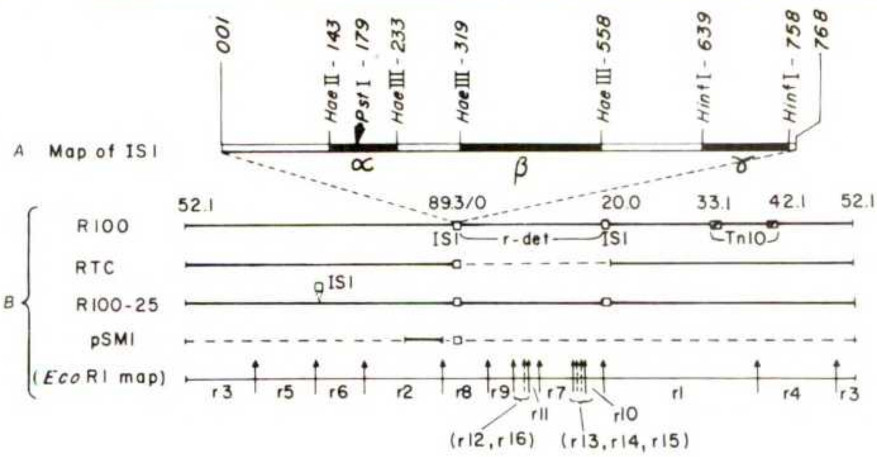


Fig. 1 *A*, Map of IS1, showing the locations of fragments used as hybridization probes. Italicized numbers indicate the coordinates of restriction endonuclease cutting sites, in nucleotides, from an end of IS1 (see ref. 6). The restriction endonuclease *Pst*I cuts IS1 at one site as indicated. *B*, Physical structures of the resistance plasmid R100 and its derivatives determined by the electron microscope heteroduplex methods[5,12,20]. All the molecules are actually circular duplexes, but the structures are displayed in a linear representation by cutting the circle at a point that lacks interesting features. All the numbers are distances in kilobases from a selected origin. Note that R100 carries two copies of IS1 (shown by open boxes) which flank the r-determinant region (r-det) containing genes responsible for resistance to chloramphenicol, sulphanilamide and streptomycin. R100 produces 16 fragments (named r with numbers) by *Eco*RI digestion, the locations of which have been determined as indicated at the bottom of the figure. The fragments, r1 and r8, contain IS1 sequences (see the gel pattern of some of the fragments shown in Fig. 2*A*). Plasmid RTC deletes the r-determinant region of R100 by recombination between the two IS1 sequences and therefore carries a single copy of IS1. This IS1 is carried by an *Eco*RI fragment of RTC, called r1′, which is a new fragment almost identical to r1. R100-25 has an additional insertion of IS1 which appears in the *Eco*RI fragment, r6, or R100 to give a new larger fragment, r6′ (see also Fig. 2*A*). pSM1 is a plasmid which deletes most of R100 but carries a copy of IS1. This plasmid was used for the isolation of the restriction fragments which are derived from inside IS1 and are used for the hybridization study.

From the nucleotide sequence of IS1 (ref. 6) the locations within the sequence of cleavage sites for restriction endonucleases such as *Hae*II, *Hae*III, *Pst*I and *Hin*fI were localized as shown in Fig. 1*A*. The DNA fragments, named α, β and γ, obtained by cleavage with any one or two of these enzymes, can be isolated from a small plasmid called pSM1, which contains one copy of IS1 (refs 6, 9), as shown in Fig. 1*A*. The 5′ end of both strands of these fragments was labelled using [γ-^{32}P]ATP and polynucleotide kinase, and the two single strands of the DNA fragment were separated as described by Maxam and Gilbert[17]. We examined whether the ^{32}P-end-labelled strands could hybridize specifically with restriction fragments containing IS1 sequences, by using plasmid DNAs such as RTC, R100 and R100-25, which are known to carry one, two and three copies of IS1, respectively. Figure 1*B* represents the physical structures of these plasmids, showing the locations of IS1 and the *Eco*RI restriction endonuclease cleavage sites. Figure 2*A* (*a–c*) shows an ethidium bromide-stained agarose gel

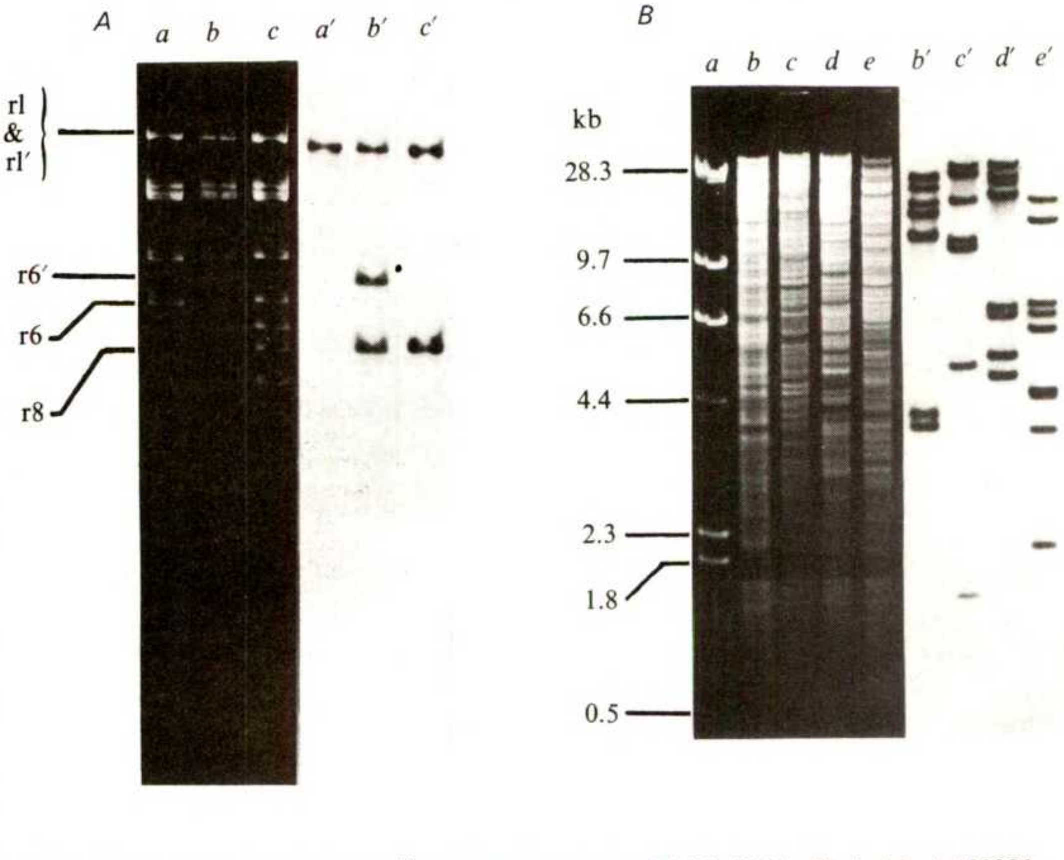


Fig. 2 Ethidium bromide-stained 0.7% agarose gels and autoradiograms of the filter obtained from the gels after hybridization with a [5′-^{32}P]-labelled IS1 DNA probe. *A*, Columns *a*, *b* and *c* are original gels of *Eco*RI digests (~1 μg) of plasmids, RTC, R100-25 and R100, respectively. r1, r6 and r8 are the fragments of R100; r1′ and r6′ are fragments unique to RTC and R100-25, respectively. r1, r8, r1′ and r6′ hybridized to the IS1 probe (β) as shown in columns *a*′, *b*′ and *c*′. *B*, Columns *b–e* are the total DNA (3 μg) of an *E. coli* K12 derivative, JE5519, digested with *Eco*RI, *Hin*dIII, *Bam*HI and *Pst*I, respectively. Column *a* is a *Hin*dIII digest of λ DNA used as length standards (in kilobases (kb))[30]. *b*′–*e*′ Are autoradiograms made after hybridization with the IS1 probe (β). For hybridization, the DNA in the gels was denatured and transferred to a sheet of nitrocellulose filter (Millipore) according to Southern[18]. The filter was then rinsed with 2 × SSC for 20 min, baked at 80 °C for 3 h and wetted from one side with 20 μl cm^{-2} of hybridization medium (50% deionized formamide and 4 × SSC) containing [5′-^{32}P]-labelled IS1 DNA (1.5×10^4 c.p.m. ml^{-1}). Wetted filters were wrapped with Parafilm and incubated in a sealed bag at 42 °C for 18–20 h. After hybridization, the filters were washed twice with 250 ml each of hybridization medium at 42 °C for 1.5 h and four times with 250 ml each of 2 × SSC at room temperature for 10 min, and blotted dry. Kodak X-ray film (SB-5) was used for autoradiography.

of the *Eco*RI digests of the plasmid DNA. The DNA fragments seen in the gel were denatured with alkali and then transferred on to a nitrocellulose filter using the method described by Southern[18]. Figure 2*A* (*a′–c′*) shows an autoradiogram of the filter obtained from the gel shown in Fig. 2*A* (*a–c*) after hybridization with the ^{32}P-labelled probe. The result shows that the *Eco*RI fragments of the plasmids which hybridize with the probe are those containing the IS1 sequences. Any one of the ^{32}P-labelled strands prepared from the three fragments described above gave results identical to those shown in Fig. 2*A* (*a′–c′*).

Table 1 Summary of number of IS1 copies in bacterial strains examined

Strain	Source* (ref.)	No. of IS1 copies
Enterobacteriaceae family		
Escherichieae tribe		
Escherichia coli		
K12: w.t. (CGSC5073)	CGSC[19]	8
K12: EMG2 (CGSC4401)	CGSC[19]	6
K12: Ymel (CGSC5032)	CGSC[19]	6
K12: W1485 (CGSC5024)	CGSC[19]	6
K12: W2637 (CGSC5290)	CGSC[19]	8
K12: W3110 (CGSC4474)	CGSC[19]	7
K12: W3110	E. Ohtsubo	8
K12: W3110 *trpE*$_{9914}$	M. Riley	9
K12: W2252 HfrC	E. Ohtsubo	7
K12: JE5519	Y. Hirota[29]	10
C	M. Riley	3
W	H. Vogel	0
Shigella dysenteriae: SH16	J. L. Rosner	>40†
sonnei	M. Pavlova	>30†
boydii	J. E. Davies	2
flexneri	J. E. Davies	>40†
Salmonella typhimurium	K. Nakamura[29]	0
Citrobacter freundii:		
ATCC8090	ATCC[29]	0
Edwardsiella tarda:		
ATCC15947	ATCC[29]	0
Klebsielleae tribe		
Enterobacter aerogenes:		
ATCC13048	ATCC[29]	0
Klebsiella aerogenes:		
T17R1	C. Yanofski[27]	1
Serratia marcescens	K. Nakamura[29]	2
Erwinieae tribe		
Erwinia amylovora:		
ATCC15580	ATCC[29]	0
Proteeae tribe		
Proteus mirabilis	K. Nakamura[27]	0
morganii	A. C. Wilson[29]	0
Gram-negative bacteria other than Enterobacteriaceae family		
Pseudomonas aeruginosa	K. Nakamura[27]	0
Acinetobacter sp. HO1-N	W. R. Finnerty[29]	0
Caulobacter crescentus CB15	A. Newton[29]	0
Myxoccus xanthus	D. R. Zusman[29]	0
Bactillaceae (Gram-positive bacterium)		
Bacillus subtilis	H. Ohtsubo	0

Species in the family Enterobacteriaceae are grouped into tribes as described in ref. 25. Bacteria in the tribe Escherichieae are more closely related to those in the tribe Klebsielleae than are the other tribes Erwinieae and Proteeae. Within the tribe Escherichieae, *E. coli* species are more closely related to *Shigella* species than are others[21–23]. The pedigree of *E. coli* K12 strains is shown in Fig. 3. The number of IS1 copies in each bacterial strain was estimated from the number of bands seen after hybridization with ^{32}P-labelled IS1 probes to the DNA, which was digested with the restriction endonuclease *Pst*I except in the cases of *S. dysenteriae* and *S. sonnei*, where *Eco*RI or *Hin*dIII was used. w.t., Wild type.

* CGSC; *E. coli* genetic stock centre at Yale University: ATCC; American Type Culture Collection.

† The exact IS1 copy number was difficult to estimate because several of the bands appeared as doublets or more.

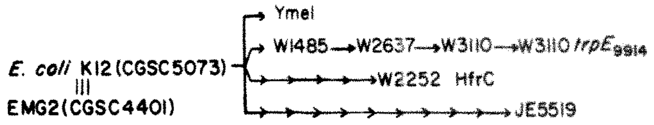


Fig. 3 Pedigree of *E. coli* K12 strains examined. Arrows indicate the number of steps required to introduce genetic markers according to Backman *et al.*[19]; *E. coli* K12 (CGSC5073) and EMG2 (CGSC4401) are supposed to be identical strains.

To determine the number of IS1 copies carried on the chromosomal DNA of *Escherichia coli* K12, total DNA was isolated from the *E. coli* derivative JE5519. Figure 2*B* shows an ethidium bromide-stained gel of the DNA cleaved with *Eco*RI, *Bam*HI, *Hin*dIII and *Pst*I, and an autoradiogram of the filter obtained from the gel after hybridization with a [5′-^{32}P]-labelled IS1 DNA probe. The results show that each enzyme generates distinctive DNA bands hybridizable with the probe. *Pst*I, which cleaves the IS1 sequence once (see Fig. 1*A*), generated the largest number of bands, 10; this is the most realistic estimate of the number of IS1 copies in the chromosome, as *Pst*I will disrupt the linkage of two copies of IS1 if present in the chromosome.

We examined the DNA from *E. coli* K12 derivatives other than JE5519. These derivatives are listed in Table 1 and their pedigree is shown in Fig. 3. Among the derivatives examined, the cells of wild-type *E. coli* K12, W1485 and EMG2 carried an F plasmid. The F plasmid DNA isolated from these strains contained no sequences hybridizable to the IS1 probes (data not shown); therefore, the difference in the number of IS1 copies in these strains is due only to the presence of IS1 in the chromosome. The results of an autoradiogram for some strains (Fig. 4) shows that *Pst*I digests of the DNA from various *E. coli* K12 strains generate not only common bands hybridizable to IS1, but also some bands unique to each strain. For example, a wild-type *E. coli* K12 strain (CGSC5073) shown in Fig. 4 was found to contain eight copies of IS1, whereas the strain EMG2 (CGSC4401), which is also supposed to be a wild-type *E. coli* K12 (ref. 19), contains six IS1 copies and shows the same pattern as that seen in Ymel and W1485. The fact that EMG2, Ymel and W1485 showed the identical pattern of six bands hybridizable with the IS1 probe suggests that this number of copies of IS1 was present in the chromosome of *E. coli* K12 when it was originally isolated from a natural source. All other derivatives contained increased numbers of IS1, demonstrated by the addition of new bands to those seen in EMG2, Ymel and W1485 (see W2637, W3110 *trpE*$_{9914}$, JE5519 in Fig. 4). It is interesting that some bacterial strains, which were identical but kept in different laboratories, contain a different number of copies of IS1.

As shown in Fig. 4, *E. coli* C contains three copies of IS1, whereas *E. coli* W contains no IS1s. This indicates that the distribution of IS1 is unique for each *E. coli* species.

We examined the DNA from various bacteria other than *E. coli*, primarily from the Enterobacteriaceae family (Table 1). As summarized in Table 1, *Klebsiella aerogenes* and *Serratia marcescens* generated one and two DNA bands hybridizable with a [5′-^{32}P]-labelled IS1 probe, respectively, but *Shigella dysenteriae* generated at least 40 bands hybridizable with the probe. None of the other strains examined generated any fragments hybridizable to the IS1 probe. The high copy number of IS1 seen in *S. dysenteriae* DNA is present not only in this species, but also in *Shigella sonnei* and *Shigella flexneri*, although the gel patterns obtained were different from each other. *Shigella boydii*, however, was found to contain only two copies of IS1.

S. dysenteriae SH16 harbours two plasmids 2 and 30 kilobases in length[20]. Hybridization analysis of these plasmids showed that the 30-kilobase plasmid contains one copy of IS1. In *S. sonnei* at least four different plasmids, about 4–7 kilobases long, were present but they contained no copies of IS1. These results indicate that the IS1 sequence seen in these two *Shigella* species are mostly present in the bacterial chromosome. We have not examined whether the other bacterial species harboured plasmids or whether these plasmids, if present, carry IS1 sequences.

A phylogenetic tree of the Enterobactericeae has recently been postulated based on immunological comparison of enzymes[21], comparison of 5S rRNA sequences[22] and comparison of the ribosomal proteins[23] (see Table 1 legend). These relationships correlate well with various other measures of biochemical relatedness, including bulk DNA homology[24], and with the numerical taxonomic groupings[25] of the Enterobacteriaceae. The distribution of IS1 sequences in various bacteria as revealed in the present analysis, however, does not correlate with the taxonomic or phylogenetic relationship of these bacteria. Besides the *E. coli* and *Shigella* strains, we detected the IS1 sequence only in *K. aerogenes* and *S. marcescens. Salmonella typhimurium, Citrobacter freundii* and *Edwardsiella tarda*, which did not contain IS1, are more closely related to *E. coli* or *Shigella* than are *K. aerogenes* or *S. marcescens* (see Table 1 legend). Although *E. aerogenes* belongs to the same tribe, Klebsielleae, as *K. aerogenes* and *S. marcescens*, it did not contain IS1. Furthermore, as described previously, *E. coli* W did not contain IS1 even though all other *E. coli* species examined did so. These results suggest that the lateral transfer of the IS1 element from one bacterial strain to other bacterial strains occurred after these bacteria had diverged phylogenetically. Genetic exchange is known to occur among Enterobacteriaceae using various systems, such as conjugation mediated by plasmids, transduction mediated by bacteriophage and direct DNA transformation. In fact, the resistance plasmids R100, R1 and R6, which are originally from *Shigella*[26] and contain two copies of IS1 (ref. 12), have been transferred to *E. coli* K12 and other bacterial strains by conjugation; the bacteriophage, P1, which can grow in *E. coli* K12 and *Shigella*, contains the IS1 sequence[27] and has been used for genetic manipulation in the *E. coli* K12 system. It is difficult, however, to determine where IS1 originally appeared.

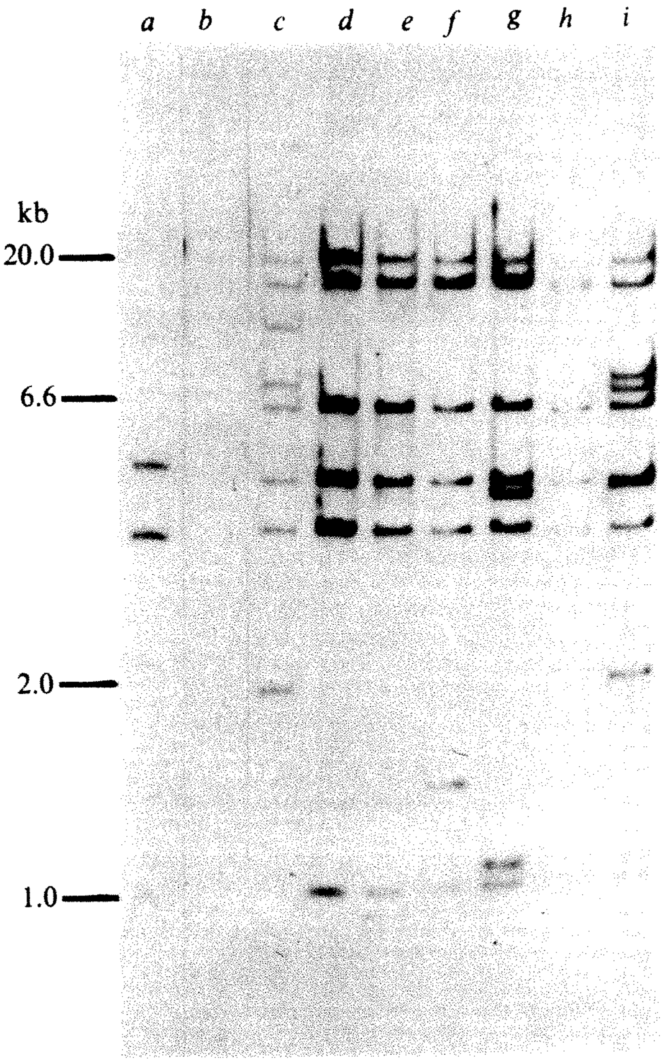


Fig. 4 Autoradiograms after hybridization of a [5'-^{32}P]-labelled IS1 fragment (γ) with *Pst*I fragments of total DNA from different *E. coli* strains: *a, E. coli* C; *b, E. coli* W; *c, E. coli* K12 (CGSC5073); *d, E. coli* K12:Ymel; *e, E. coli* K12:W1485; *f, E. coli* K12:W2637; *g, E. coli* K12:W3110 *trpE*$_{9914}$; *h, E. coli* K12:W3110 (CGSC4474); *i, E. coli* K12:JE5519. *Hin*dIII digests of phage λ DNA were used for calibrating fragment sizes. kb, Kilobases.

It is known that some transposable DNA elements containing resistance genes for antibiotics and enterotoxin genes are flanked by two IS1 sequences[12,14,16,28]. It is known that many resistant bacteria have different resistance spectra due to the presence of various kinds of plasmids, a fact first reported in *Shigella* strains in Japan[26]. This suggests that the elaborate recombination events responsible for the generation of resistance plasmids may have been facilitated by the large number of IS1 sequences in *Shigella*.

In addition to IS1 and several insertion sequences including IS2 and IS3 (see ref. 3), there is a group of translocatable elements responsible for resistance to various antibiotics. The investigation of the presence of these elements will also be important for the elucidation of the origin of resistance plasmids in bacteria.

We thank Monica Riley and Algis Anilionis for the gift of bacterial strains and bacterial DNA, and Jonathan Rosen and Tom Ryder for critical reading of the manuscript. K.N. was supported under the USPHS grant GM19043 (Principal Investigator, M. Inouye). This work was supported by USPHS research grants GM22007 to E.O. and GM26779 to H.O.

Received 23 September; accepted 27 November 1980.

1. Starlinger, P. & Saedler, H. *Curr. Topics Microbiol. Immun.* **75,** 111–152 (1976).
2. Nevers, P. & Saedler, H. *Nature* **268,** 109–115 (1977).
3. Campbell, A. *et al.* in *DNA Insertion Elements, Plasmids, and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S.) 15–22 (Cold Spring Harbor Laboratory, New York, 1977).
4. Fiandt, M., Szybalski, W. & Malamy, M. H. *Molec. gen. Genet.* **119,** 223–231 (1972).
5. Hirsch, H-J., Starlinger, P. & Brachet, P. *Molec. gen. Genet.* **119,** 191–206 (1972).
6. Ohtsubo, H. & Ohtsubo, E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 615–619 (1978).
7. Johnsrud, L. *Molec. gen. Genet.* **169,** 213–218 (1979).
8. Johnsrud, L., Calos, M. P. & Miller, J. H. *Cell* **15,** 1209–1219 (1978).
9. Ohtsubo, E., Rosenbloom, M., Schrempf, H., Goebel, W. & Rosen, J. *Molec. gen. Genet.* **159**, 131–141 (1978).
10. Ohtsubo, E., Zenilman, M. & Ohtsubo, H. *Proc. natn. Acad. Sci. U.S.A.* **77,** 750–754 (1980).
11. Iida, S. & Arber, W. *Molec. gen. Genet.* **153,** 259–269 (1977).
12. Hu, S., Ohtsubo, E., Davidson, N. & Saedler, H. *J. Bact.* **122,** 764–775 (1975).
13. Gottesman, M. M. & Rosner, J. L. *Proc. natn. Acad. Sci. U.S.A.* **72,** 5041–5045 (1975).
14. MacHattie, L. A. & Jackowski, J. B. in *DNA Insertion Elements, Plasmids, and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S.) 219–228 (Cold Spring Harbor Laboratory, New York, 1977).
15. MacHattie, L. A. & Shapiro, J. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1490–1494 (1978).
16. So, M., Heffron, F. & McCarthy, B. J. *Nature* **277,** 453–456 (1979).
17. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
18. Southern, E. *J. molec. Biol.* **98,** 503–517 (1975).
19. Bachmann, B. J. *Bact. Rev.* **36,** 525–557 (1972).
20. Ohtsubo, H. & Ohtsubo, E. in *DNA Insertion Elements, Plasmids, and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S.) 49–63 (Cold Spring Harbor Laboratory, New York, 1977).
21. Cocks, G. T. & Wilson, A. C. *J. Bact.* **110,** 793–802 (1972).
22. Hori, H. *Molec. gen. Genet.* **145,** 119–123 (1976).
23. Hori, H. & Osawa, S. *J. Bact.* **133,** 1089–1095 (1978).
24. Sanderson, K. E. *A. Rev. Microbiol.* **30,** 327–349 (1976).
25. Buchanan, R. E. & Gibbons, N. E. (eds) in *Bergey's Manual of Determinative Bacteriology* 8th edn (Williams & Wilkins, Baltimore, 1974).
26. Mitsuhashi, S. (ed.) in *Transferable Drug Resistance Factor R* (University of Tokyo Press, 1971).
27. Arber, W. *et al. Cold Spring Harb. Symp. quant. Biol.* **43,** 1197–1208 (1979).
28. Meyer, J. & Iida, S. *Molec. gen. Genet.* **177,** 261–270 (1980).
29. Nakamura, K., Pirtle, R. M. & Inouye, M. *J. Bact.* **137,** 595–604 (1979).
30. Robinson, L. H. & Landy, A. *Gene* **2,** 1–31 (1977).

BOOK REVIEWS

Cognition at the crossroads

John C. Marshall

Language and Learning: The Debate Between Jean Piaget and Noam Chomsky. Edited by M. Piattelli-Palmarini. Pp.409. (Harvard University Press/Routledge and Kegan Paul: 1980.) $20, £9.75.

IF THIS book did not exist, only a supreme ironist could have invented it.

In 1975 the Royaumont Center for a Science of Man revived an ancient tradition: before a vociferous audience of highly distinguished molecular biologists, anthropologists, mathematicians, psychologists, computer scientists and philosophers, two pre-eminent scholars, Jean Piaget and Noam Chomsky, conducted a disputation on the (biological) nature of the human soul. The only aspect of the debate which is beyond dispute is that the central figures are indeed the two scholars who have had the greatest impact upon the psychological sciences in the second half of this century.

It is ironic, but not altogether surprising, that neither man is a psychologist by training. Piaget began his career in biology and then spent some 50 years observing the development of reasoning in children, much as a naturalist might; Chomsky studied logic, philosophy of science and structural linguistics before noticing, as a graduate student, that the grammars he was writing could be interpreted as theories of language-acquisition.

The second irony, or at least shock, arises from being confronted with such diametrically opposed responses to the command "Etonnez-moi!". Piaget has astonished us with data which, if valid, suggest that children live in a world which is incommensurate with that of their parents. It is not that children merely know fewer facts than we do, but that their entire interpretive framework, their logic, differs from ours. For Piaget, the child believes that water poured from a short, wide glass into a tall narrow one increases in volume; likewise, the child may grant that A is larger than B, and that B is larger than C, but will not endorse the inference that therefore A *must* be larger than C. For the child, the latter claim is an empirical issue to be settled by observation not speculation. Critics of Piaget typically attack the database from which his startling conclusions are drawn: the child misunderstood the question, became confused by the perceptual salience of the array, or forgot the premises before getting around to the answer. As Jacques Mehler points out, the experimentalists' standard retort to Piaget is thus not to explain his results but rather to explain them away.

Camera Press

Jean Piaget — children think differently

Contrariwise, the primary data that Chomsky has drawn attention to are almost trivially plain and unimpeachable. For example: the sentence "We like each other" is, roughly, a paraphrase of "Each of us likes the other." But "We expect John to like each other" does not mean "Each of us expects John to like the other"; indeed the former string is not even a well-formed sentence of English. The facts here and in myriad other cases are not in doubt. Chomsky astonished us by demanding an explanation for such obvious facts. That is, he outlined a program for linguistic theory: to write with a finite number of rules, some of them recursive, a grammar that generated all the sentences of a language and none of the non-sentences; to formulate the rules in such a way that they automatically assign the correct structural description to each generated sentence; to discover the most restricted schematism for rule-types that will allow the description of any natural language while formally excluding rule-types that are found in no natural language. The very idea of such a program was breathtaking and even the admittedly limited success with which Chomsky and other linguists have been able to carry it out is a good deal more impressive than anything else that has been achieved in the human sciences. Chomsky's revolutionary goals rapidly came to define the field and, except in a few benighted backwaters, linguistics now is what Chomsky in 1955 said it should be. "Internal" critiques of Chomsky therefore take just the form one would expect: particular rules and rule-schema are attacked because they overgenerate, undergenerate or otherwise assign incorrect structural descriptions.

The third irony lies in the topic of debate. Piaget has rarely, if ever, studied language. Although the title of his first major publication contains the word "language", Piaget's concern was not how children talk but rather what they talk about. In the opinion of many psychologists, particularly those who follow young children around with tape-recorders, Chomsky has never studied language-acquisition either. Finally, neither Piaget nor Chomsky consider that learning *as classically conceived* plays more than a minor role in the acquisition of any significant body of knowledge.

One might predict then that the Royaumont Center has devised a recipe for disaster, but happily this is not so. All the participants rise magnificently to the occasion and a display of truly virtuoso editing from Massimo Piattelli-Palmarini has converted the proceedings into a first-rate book.

Chomsky's position has a stark simplicity. The adult's linguistic competence can be represented by a generative grammar that pairs sound and meaning correspondences across the infinite number of sentences of the language. This grammar is an abstract description of one component of mind, a theory of the steady-state attained. Universal grammar, by contrast, characterizes the properties that languages hold in common or that constrain the space wherein they may vary. These principles, along with parameters that allow for variation between languages, are attributed to the child's initial state; they form boundary conditions within which the learner attempts to locate the specific grammar of the particular language to which he is exposed. The function of exposure to a particular linguistic environment is to set the parameters left open in universal grammar; does the

Noam Chomsky — universal grammar

language have structure-dependent movement rules (English) or not (Japanese)? Is the order of base elements subject–verb–object (English) or subject–object–verb (German)? Can tensed sentences be overtly subjectless (Italian) or not (English)? Once the value of a parameter has been fixed, the relevant structures then interact with universal principles so that a range of other properties of the language follows deductively. It is in this sense that the child comes to know more than he experiences.

Chomsky's program ascribes to the genotype some highly constrained, domain-specific principles. Thus the pattern of the linguistic examples quoted earlier is interpreted by postulating a condition on rules of sentence grammar such that if an embedded sentence has a subject (*John* in "We expect John to like each other"), only that subject is accessible to the rule of reciprocal interpretation. The effect of the specified subject condition is to force the complement of the ungrammatical string into the anomalous structure "John likes each other". Many of the participants find the imputation outrageous. They concede that radical behaviourism is dead (William of Occam never intended psychologists to cut their throats with his razor) but baulk at attributing the specified subject condition to the genome. Surely general principles of self-regulation, sensori-motor intelligence or learning-strategies will suffice for the child to acquire his language? Each time that one of these suggestions crops up, Chomsky asks for an example of a general learning strategy (or whatever) that would lead to the adult's language obeying the specified subject condition. And each time he does so, the silence is deafening.

Mention of learning inspires Jerry Fodor to essay his now notorious argument that all knowledge is innate (*The Language of Thought*; Harvester, 1975). Fodor observes that the only accounts of "concept learning" ever developed are variants of the theory of inductive inference. A hypothesis ("All emeralds are green") is projected, and if confirmed (= not disconfirmed) by the evidence, belief is fixed. The theory entails that "the concepts that figure in the hypotheses that you come to accept are not only *potentially* accessible to you, but are actually exploited *to mediate the learning*". Inductive inference is normally thought of as the cornerstone of empiricist theories of learning, but, as Fodor concludes, the model "assumes the most radical possible nativism: namely, that any engagement of the organism with its environment has to be mediated by the availability of any concept that can eventually turn up in the organism's beliefs". The upshot is to make innate all the semantic primitives or meaning postulates that characterize every concept from elephants to electrons. They lurk inside the organism waiting to be triggered by the environment. Once more, many of the participants at the symposium are affronted although the ghost of Socrates can be heard cheering from the sidelines. Fodor's argument has circulated widely for five years, and is generally regarded as misconceived. Yet no counter-argument is known nor any convincing alternative account of learning whereby the conclusion can be avoided. Conceptual *development* — the acquisition of qualitatively stronger logics, in Piaget's terms — is not ruled out by Fodor's account, but it must arise from maturation, not "learning".

The metaphor of the growth of mental organs is congenial to both Chomsky and Piaget, though Chomsky takes the stricter line. Piaget has long held the view that an organism can "assimilate" certain structural properties of the environment, thereby allowing learning *sine* learning. The notion has affinities with J. J. Gibson's theory of "direct perception". Piaget's "constructivism" attempts to show how "new operations and structures" can arise from interaction with the environment (and hence undermine Fodor's conclusion). A favourite example of Piaget's is the concept of "reversibility". The idea originates, Piaget believes, when the child notices that he can, for example, pick up a ball, replace it and pick it up again. Expressed first through the operation of "sensori-motor intelligence", the concept of reversibility eventually becomes available for incorporation into more abstract schemata (if $A + B = C$, then $C - B = A$). Piaget, however, fails to show how the concept could arise from sensori-motor play or become generalized *except* in an organism that was specifically pre-adapted to encode its experience in terms of reversibility. We thus return full-circle to Fodor's point.

Piaget tries to reinforce his position by adopting an analogy that the molecular biologists find objectionable. He employs a quasi-Lamarckian account of phenocopies in which the environment can *change* the structure of an organism. This holds little appeal for François Jacob, Antoine Danchin, Jean-Pierre Changeux and Jacques Monod. Jacob stoutly defends neo-Darwinism: phenocopies constitute "only a slight variation within the realm of what the genotype allows"; Danchin stresses how molecular biology has advanced "*without allowing the intervention of even the least instructive notion on the part of the environment*"; and Changeux expounds his elegant theory of synaptic stabilization in which the effect of the sensory environment is simply to select a pattern of connectivity by pruning unused elements.

As Piattelli-Palmarini emphasizes in his superb introduction, the research program of molecular biology, conducted at the level of concrete mechanism, bears a striking resemblance to Chomsky's program, although the latter enterprise only characterizes the "abstract conditions that unknown mechanisms must meet" (Chomsky). Tracing these conditions back to their underlying mechanisms should keep us all busy for many a year. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at The Radcliffe Infirmary, Oxford.

Pions and nuclei

D. V. Bugg

Theory of Meson Interactions with Nuclei. By Judah M. Eisenberg and Daniel S. Koltun. Pp.403. (Wiley-Interscience: 1980.) £21.25, $49.95.

NUCLEAR forces are due to exchange of mesons such as the pion. If, in nuclear structure physics at low energy, one gives nucleons a nudge weak compared with these strong forces, a rich excitation spectrum results. The meson factories now allow systematic study of much larger excitations, using beams of pions or nucleons. The scattering is dominated by nucleon excitations such as the Δ (1230) at 300 MeV in the πN system. To make sense of the data, it is necessary (i) to know the fundamental interactions (πN, NN and $N\Delta$ including off-shell extrapolations), and (ii) to solve the many body problem. Since the interactions are strong, it is generally inadequate to approximate the problem by single scattering from individual nucleons.

This book sets out to treat the multiple scattering series in systematic fashion. The exposition is concise, elegant and reasonably transparent, and brings order to the chaos of approximations in the literature. The aim is to educate research workers and students in this immediate field, though students will need some previous exposure to formal scattering theory up to the Lippmann–Schwinger equation. Armed with this and the rather brief exposition of πN dynamics given in the book, they can follow clear trails through the jungle of multiple scattering series (to fourth order, including nuclear excitation and anti-symmetrization), Fadeev equations, dispersion theory, optical models and Glauber approximation to confront data on elastic and inelastic nuclear scattering, knockout reactions, pion absorption and mesic atoms. The formal theoretical machinery is all there, and the precise relationship between exact solutions and approximations is quite clear.

The disappointment of the book is that, after all this effort, little qualitatively new physics emerges. This reflects on the subject, not the presentation. As yet, one does not know whether failures of calculation arise from approximations (necessary to limit computation) or from ignorance of fundamental interactions. This is a pity, since data like these potentially hold the key to understanding Δs in nuclei and nuclear matter at high densities (hence high Fermi levels), where a phase transition to a pion condensate may occur. ☐

D. V. Bugg is Professor of Nuclear Physics at Queen Mary College, University of London, and an SRC Senior Fellow.

Towards harmony in Polar geobotany

Stanwyn G. Shetler

The Arctic and Antarctic: Their Division into Geobotanical Areas. By V.D. Aleksandrova. Pp.247. (Cambridge University Press: 1980.) £15, $34.50.

AT LAST an attempt has been made at reviewing the literature on the vegetation cover of both the Arctic and the Antarctic and also at classifying all Polar geobotanical areas from a single point of view. Such an extraordinary synthesis is the more important for coming from the eminent Soviet geobotanist, Vera Aleksandrova. One of the world's foremost authorities, she has been a leading figure in ecology for nearly three decades in the country with the lion's share of the Arctic.

Aleksandrova's contribution, now translated from Russian, should do much to bridge the gap between Soviet and other Polar ecologists. Apart from the formidable language barrier, there are fundamentally different traditions in concept and method; even the very terms are different.

Soviet geobotanists, in the traditional European mould, continue to classify vegetation as an end, a science, in itself. This book is fine testimony to the continuing preoccupation of the Soviet school of geobotany with discovering the ultimate phylogenetic ("phylocoenogenetic") hierarchy of plant communities. Western, particularly North American, ecologists largely have taken a much more pragmatic approach. Classification is a means to an end, a means of organizing the state of knowledge of vegetation.

Aleksandrova's objectives, to subdivide and classify the geobotanical areas of Polar lands and at the same time to summarize succinctly the vast amount of data on vegetation cover, have been met well. Interesting and potentially useful as her classification is, her vegetation summary is certain to prove the more valuable to botanists generally, who until now have not had a single place to turn for such a detailed and yet generalized overview of the vegetation cover in the Soviet Arctic. Thus her bias toward greater detail and precision for Soviet territory is easily forgiven. We can be grateful that the basis did not exist for casting her study in the mould of some recondite phytosociological system of classification with its pointless nomenclature, even though she would not agree.

She divides the Arctic into the Tundra Region and Polar Desert Region. The first is further subdivided into two subregions: Arctic Tundras and Subarctic Tundras, differentiated mainly by the presence of shrubby birches (*Betula*) in the zonal associations of the latter but not the former. Both are further subdivided into provinces, subprovinces and districts. The Antarctic is divided into the Subantarctic Cushion Plant Region and the Antarctic Polar Desert Region. The latter is an analogue of the Arctic Polar Desert Region, but she finds no South Polar tundra analogue.

Her system of classification might be termed holistic for drawing on such a broad range of data. Two types of features are used: diagnostic and characterizing. The diagnostic features are zonal features that govern the delimitation of the major geobotanical areas. Ultimately, however, Aleksandrova, like geobotanists in general, must place strong reliance on floristics, little else being as tangible and objective. Certain species (*Salix polaris*, for example) do yeoman service in the characterizations.

Despite its small size, the book is rich in detail, perhaps too much so in places. This detail and the liberal use of phytosociological jargon do not make for light reading; it is clearly a book to refer to, not to read straight through. The book would have benefited enormously from some well-chosen photographs of the main types of vegetation discussed. A bit too much geography, especially of the USSR, is taken for granted, and the reader wishes for a good map of the Arctic with major place names shown. A glossary of terms used in Soviet ecology would also be most useful.

The book is well produced with a minimum of errors, all minor. The extensive bibliography is invaluable. With few exceptions it seems amazingly complete for North America (Raup's work on the Mackenzie Territory seems to have escaped review). As translator, Doris Löve has done a remarkable job, and she is to be commended for undertaking this task on behalf of all of us.

Criticisms notwithstanding, this is an outstanding piece of work that will go far toward fostering understanding and harmony in Polar geobotanical research, while at the same time laying foundations for new efforts. This instantly indispensable reference is sure to become a classic. ☐

Stanwyn G. Shetler is Associate Curator in the Department of Botany, Smithsonian Institution, Washington DC.

TAS, 45 years on

M.J. Seaton

Atomic Structure. By E.U. Condon and Halis Odabaşi. Pp.658. (Cambridge University Press: 1980.) Hbk £30, $83.50; pbk £9.95, $26.

The Theory of Atomic Spectra by Condon and Shortley, published in 1935 and familiarly known as TAS, is one of the great scholarly classics of twentieth-century scientific literature; it was reprinted many times by Cambridge University Press but there was never a second edition. From 1962, when he went to the University of Colorado, until near the time of his death in 1974, Ed Condon worked on a new book in collaboration with Halis Odabaşi who is now at Boğaziçi University, Turkey. I witnessed the work in progress on the tenth floor of the JILA tower; its fruits have now been prepared for publication by Odabaşi with assistance from many people, particularly John Cooper and Roy Garstang, and published as *Atomic Structure*.

We can cherish the treasures in an unfinished symphony without concealing the shortcomings of the work as a whole. For me the main source of delight in this new book is the first chapter, concerned with the historical developments which culminated in the modern formulation of quantum mechanics. Others may find particular merit in the chapters on group theory (of this subject, it was said in the introduction to TAS that "We manage to get along without it"). *Atomic Structure* is concerned exclusively with the non-relativistic theory, although, in his preface, Odabaşi tells us that he and Garstang are working on another volume, intended as a sequel, in which the relativistic theory will be discussed. The main new developments in non-relativistic atomic structure physics since TAS was published are the introduction of Racah algebra techniques, the carrying out of much more accurate calculations using computers and the accumulation of a wealth of new experimental data. Of these, a chapter is devoted to the first but the other two are sadly neglected. Much of the book is concerned with the central-field model, in particular the calculation of the algebraic coefficients which occur in the diagonal matrix elements of the Hamiltonian using a single-configuration representation. Having obtained such coefficients, Condon and Shortley in TAS immediately consider whether "the formulas agree with the data in any sense". Their first confrontation is for the configuration p^2, for which the single-configuration theory tells us that there are three terms, ^3P, ^1D and ^1S, and gives us a predicted ratio of separations

$$[E(^1S)-E(^1D)]/[E(^1D)-E(^3P)] = 1.5$$

this ratio being determined entirely by the algebraic coefficients and independent of the choice of orbital functions. The comparison with experimental ratios given in TAS, 1.13 for C I, 1.14 for N II, 1.14 for O III and so on, shows the inherent limitations of the single-configuration theory, which may give useful information on the relative positions of the terms but fails to give precise quantitative information on energy levels and other properties of atoms. While many similar comparisons were given in TAS, only a limited number are given in *Atomic Structure*, and these only in a concluding chapter and in the context of Hartree–Fock theory. The inherent limitations of the single-configuration theory are thus obscured, and a large amount of more recent work, using multi-configuration methods, is neglected completely. (In such work, the entire problem is handled on a computer — the algebra, the calculation of radial functions, the diagonalization of the Hamiltonian matrix and the computation of matrix elements required for the determination of other atomic properties.)

Advances in studies of atomic spectra and atomic structure which have taken place during the past 45 years are such that it becomes increasingly difficult to describe them in a single comprehensive book. A number of specialist works have been published, and doubtless others will appear, but the need for a more comprehensive work remains. Although *Atomic Structure* fails to meet this need satisfactorily, it contains much that is of value. Professor Odabaşi, others who have assisted and Cambridge University Press should be thanked and congratulated for having made it available to us. □

M. J. Seaton is Professor of Physics in the University of London, and Fellow-Adjoint of the Joint Institute for Laboratory Astrophysics (JILA), University of Colorado, and National Bureau of Standards.

Tool use — fundamentals and philosophy

Toshisada Nishida

Animal Tool Behavior: The Use and Manufacture of Tools by Animals. By Benjamin B. Beck. Pp.307. (Garland: New York, 1980.) $24.50.

THE biggest difficulty in studying tool-using behaviour is deciding what constitutes tool use and what does not — no two scientists would completely agree with any definition. Beck illustrates this fundamental problem right at the beginning of his book in providing a test for the reader. He presents 50 animal behaviours from which those representing tool use must be chosen — I only just passed, answering 35 correctly.

The author goes on to catalogue, in considerable detail, observed tool behaviour, and provides a useful classification of tool-manufacture into four components: *detach, subtract, add or combine* and *reshape*. There are a few omissions and some examples with which I would disagree — fishing probes made of the mid-rib of leaves may be an overlooked example of *subtract* and chimpanzees do not "dip" but rather fish for *Camponotus* ants. Beck then reconsiders his catalogue according to the major functions of tool use: *to extend the user's reach, to amplify mechanical force, to augment agonistic display* and *to allow more effective control of fluids*. Again this is a useful categorization, but I would prefer to substitute "augment signal value of display" for "augment agonistic display" as chimpanzees have been seen to use tools even in courtship and other non-aggressive displays.

How do animals begin to use tools? Beck stresses that the discovery of novel tool-using patterns could result fortuitously from exploratory object manipulation and reviews the factors that influence the development of tool use: maturation, and associative and early experience. He then goes on to criticize Alcock's view that tool-users are morphologically ill-adapted to compete for resources and that tool-use compensates for lack of biological equipment. Since all behaviour can be viewed as compensating for lack of morphological structure, the author argues that Alcock's tenet has no special significance for the evolution of tool-using behaviour. Taking chimpanzees as an example, he also persuasively argues against another thesis of Alcock's: that tool-users have invaded niches that are not at all characteristic of their phylogenetic group.

In considering the cognitive aspects of tool use, Beck questions whether the behaviour can be called "intelligent". Pointing out that non-tool behaviour of herring gulls (shell-dropping) and chimpanzees' tool behaviour (termiting) share cognitive similarites, he attacks the notion that tool behaviour necessitates a cognitive sophistication which is unique and more advanced than that subserving non-tool behaviour. He concludes that there is no unique, orderly relationship between tool use and intelligence, and that social and feeding behaviours are probably far more important factors in the evolution of intelligence. Similar arguments are put forward in reviewing the development of human intelligence and tool use.

Finally, he turns to the philosophical implications of his work in remarking that "tool behaviour should not only be stripped of its role as a primary shaper of our own species, but should also be recast as the progenitor of the engine of our destruction". This pessimism represents a trend stemming from the author's opposition to the ethnocentrism of civilizations and the anthropocentrism of our own species, one which has perhaps been greatly influenced by his anxiety about the future of an Earth under threat from a number of human activities. I don't support all of his remarks. I think that the primary shaper of our species is socio-economic behaviour (including subsistence tool use) and that communicative behaviour is secondary. Moreover, the lack of meaningful relationships between tools and intelligence in other species does not necessarily negate its possibility in *Homo*. In this context, I do not understand why he omits Penfield and Rasmussen's famous contribution to cerebral physiology: hand areas of both motor and sensory cortex in human beings are relatively much larger than corresponding areas in apes and monkeys. This can be easily understood by assuming that tools were one of the major selective agencies in human evolution. Indeed, one of the faults of the book is the lack of a systematic review of gripping organs, grasping methods, handedness and skills in primate tool-use.

Yet this is a stimulating book, rich in suggestions, and undoubtedly the best sourcebook available on animal tool behaviour. □

Toshisada Nishida is a Lecturer in the Department of Anthropology at The University of Tokyo.

ANNOUNCEMENTS

Awards

The 1980 Smith Kline & French Foundation biennial prize for research in clinical pharmacology has been awarded to **Dr M. L'E Orme** (Liverpool University) for his work on oral contraceptive steroids.

The Fyssen Foundation is to award fellowships for the training and support of research fields as ethology, palaeontology, archaeology, anthropology, psychology, logic and the neurosciences. Priority will be given to French scientists wishing to work abroad and to foreign scientists wishing to work in French laboratories. Study grants will normally be granted for one year but may be extended up to three. Applications to the Secretariat of the Foundation, 194, rue de Rivoli, 75001 Paris, France before 1 May 1981. Nominations for the 1981 Prize of the Fyssen Foundation should be submitted to the above address before 1 September 1981.

Young people under 31 working in industry and the universities, are invited to enter for the Suttle Award of the Filtration Society. The award is to be given for the best paper on filtration/separation published between June 1979 and June 1981. Full details from: the Hon. Secretary, The Filtration Society, 2 Woodstock Road, Croydon, Surrey, UK).

Appointments

The Public Health Laboratory Service Board has appointed **Dr J. E. M Whitehead** as director of the Service in succession to Sir Robert Williams who will retire on 1 July 1981.

Dr M. Margaret Clark, has been elected president of the American Anthropological Association.

Dr Gordon Leedale, has been appointed by the University of Leeds to a personal Chair in botany.

Meetings

12-13 March, **Animal Models in Parisitology,** London (C.V. Clark, MRC Laboratory Animals Centre, Woodmansterne Rd, Carshalton, Surrey, UK).

6-12 March, **Nitrogen Recycling in Ecosystems in Latin America and the Carribean,** Colombia (T. Rosswall, SCOPE/UNEP International Nitrogen Unit, The Royal Swedish Academy of Sciences, S-104 05 Stockholm, Sweden).

17-20 March, **Guidelines for Control of Toxic and other Hazardous Wastes,** Wiesbaden (WHO, 8 Scherfigsvej, DK-2100 Copenhagen O, Denmark).

18 March, **Polar Ice and Climate,** London (Royal Meteorological Society, James Glashier House, Grenville Place, Bracknell, Berks, UK).

16-17 March, **Research on Local Wind Environments,** Cambridge (The Institute of Mathematics and its Applications, Maitlanf House, Warrior Square, Southend-on-Sea, Essex, UK).

31 March — 2 April, **Occupational Hazards of the Health Professions,** Homburg (WHO, 8 Scherfigsvej, DK-2100 Copenhagen O, Denmark).

8-13 March, **Structure and DNA–Protein Interactions of Replication Origins,** Salt Lake City, ICN–UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

8-13 March, **Genetic Variarion among Influenza Viruses,** Salt Lake City, (ICN–UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

15-20 March, **Developmental Biology using Purified Genes,** Keystone (ICN–UCLA Symposia, Molecular Biology Institute, University of California, Los Angeles, California 90024).

18-19 March. **Information Society: How to cope with the Electronic Revolution,** Rotterdam (Information Dept TNO, PO Box 297, 2501 BD The Hague, The Netherlands).

19-20 March, **Fouling of Heat Exchangers,** Birmingham (Institution of Chemical Engineers, 12 Gayfere St, London SW1, UK).

20 March, **Bracken in Scotland,** Edinburgh (The Meetings Secretary, The Royal Society of Edinburgh, 22, 24 George St, Edinburgh, UK).

23-24 March. **Ion Exchange Technology,** London (Institution of Chemical Engineers, 12 Gayfere St, London SW1, UK).

23-27 March, **Medicinal Chemistry,** Nottingham (Miss L. Hart, Royal Society of Chemistry, 30 Russell Square, London WC1, UK).

24 March, **Applications of Thermal Methods of Analyses in the Pharmaceutical Industry,** Nottingham (Dr D.J.W. Grant, Dept of Pharmacy, University Park, Nottingham, UK).

26-28 March, **Biophysics after 21 Years,** Leeds (Prof. A.C.T. North, The Astbury Dept of Biophysics, The University of Leeds, Leeds, UK).

24-27 March, **Scientific Importance of High Angular Resolution at IR and Optical Wavelengths,** München (ESO Conference, Karl-Schwarzschild-Straße 2, D-8046 Garching Bei München, FDR).

27-28 March, **Underwater Biology in The British Isles,** London (Paul Clark, Dept of Zoology, British Museum (Natural History), Cromwell Rd, London SW7, UK).

29 March-3 April, **Specialised Transmission Electron Microscopy Course,** Leeds (The Administrator, Royal Microscopical Society, 37/38 St Clements, Oxford, UK).

31 March, **Ethical and Legal Aspects of Radiological Protection,** London (Prof. J.H. Martin, Dept of Medical Biophysics, University of Dundee, UK).

1 April, **1st Annual EM Lecture,** Leeds The Administrator, Royal Microscopical Society, 37/38 St Clements, Oxford UK).

3 April, **Deep levels in III–V Semiconductors,** London (The Institute of Physics, 47 Belgrave Square, London SW1, UK.

7-8 April, **University Research — Who Benefits?,** Sheffield (The Filtration Society, 2 Woodstock Rd, Croydon, Surrey, UK).

7-8 April, **Soil Physical Conditions and Crop Growth,** London (Dr D.V. Crawford, University of Nottingham School of Agriculture, Sutton Bonnington, Loughborough, Leicester, UK).

8 April, **Calibration and Standarisation of Particle Sizing Methods,** Teddington (Dr R. Wilson, Particle Metrology Section, National Physical Laboratory, Teddington, Middlesex, UK).

8-10 April, **Formation and Evolution of the Hellenic Arc and Trench,** Athens (Xavier Le Pichon, Laboratoire de Géodynamique, Université Pierre et Marie Curie, 4 Place Jussieu, 75320 Paris Cedex 01, France).

13-14 April, **Ion Exchange,** Nottingham (Miss L. Hart, Royal Society of Chemistry, 30 Russell Square, London WC1, UK).

13-14 April, **Mechanisms of Steroid Action,** London (Mrs J. Kruger, Dept of Pharmacology, University College, Gower St, London WC1, UK).

13-16 April, **1st Meeting of the European Union of Geosciences,** Strasbourg (EUG Organizing Committee, Institut de Physique du Globe de Paris, Universite Paris VI, 4 Place Jussieu, 75320 Paris Cedex 01, France).

22-25 April, **Inflammation Markers, Lyon** (A. Roulett, Hopital Jules Courmont Sainte Eugenie, 69310 Pierre Benite, France).

26 April-1 May, **Freeze-etching in Electron Microscopy,** Woods Hole (Morton D. Maser, Marine Biological Laboratory, Woods Hole, Massachusetts 02543).

5 May, **International Symposium on Crop Protection,** Gent (Faculteit van de Landbouwwetenschappen, Coupure links, 533, B-9000 Gent, Belgium).

6-9 May, **Monoclonal Antibodies and Ultrasensitive Immunoassay in Human Diagnosis and Monitoring of Therapy,** *and* **Calciotropic Hormones Methods and Clinical Applications,** Gardone Riviera (RIA 81, Fondazione Giovanni Lorenzini, Via Monte Napoleone, 23, 20121 Milan, Italy).

10-12 May, **Biophysical Aspects of Immunology,** Galilee (Prof. Ruth Arnon, Dept of Chemical Immunology, Rehovot, Israel).

18 May, **Aspects of Xylem Differentiation,** Princes Risborough (Dr J. R. Barnett, Plant Sciences Laboratories, University of Reading, Whiteknights, Reading, Berks, UK).

18-21 May, **Space Manufacturing,** Princeton (B. Evans, Space Studies Institute, PO Box 82, Princeton, New Jersey 08540).

27-29 May, **35th Annual Frequency Control Symposium,** Philadelphia (Mrs L. Hildebrandt, US Army Electronics Research & Development Command, Fort Monmouth, New Jersey 07703).

28-31 May, **4th Greek Congress of Anaestesiology,** Athens (Greek Society of Anaesthesiologists, 34 Dragoumi St, Athens,T. 612, Greece).

1-5 June, **Phosphorous Chemistry,** North Carolina (Prof. L. D. Quin, Paul M. Gross Chemical Laboratory, Duke University, Durham, North Carolina 27706).

3-4 June, **Prospective Biological Products for Viral Diseases,** Munich (P. A. Bachmann, WHO Collaborating Centre, Koniginstrape 49, 8000 Munich 22, FRG).

9-11 June, **34th Chemists' Conference,** Scarborough (British Steel Corporation, Teesside Laboratories, Grangetown, Middlesborough, Cleveland, UK).

16-17 June, **Chromatography in Biochemistry, Medicine and Environmental Research** *and* **Mass Spectrometry in Biochemistry, Medicine and Environmental Research,** Venice (Italian Group for Mass Spectrometry in Biochemistry and Medicine, Via Eritrea, 62, 20157, Milan, Italy).

24 June-5 July, **Targetting of Drugs,** Cape Sounion Beach (Dr G. Gregoriadis, Division of Clinical Sciences, Clinical Research Centre, Watford Rd, Harrow, Middx, UK).

24-27 June, **Animal, Plant and Microbial Toxins,** Marseille (Dr H. Rochat, Laboratoire de Biochemie, Faculté de Médecine, Boulevard P. Dramard, 13326 Marseille Cedex 3, France).

29 June-3 July, **Statistical Methods in Cancer Epidemiology,** Lyon (The Head, Research Training and Liasison Programme, International Agency for Research on Cancer 150, cours Albert Thomas, 69372 Lyon Cedex 2, France).

28 June-3 July, **Biological Psychiatry,** Stockholm (C. Hamilton, Stockholm Convention Bureau, Jakobs Torg 3, S-111 52 Stockholm, Sweden).

1-4 July, **Minimal Invasive Cancer,** Graz (Interconvention, PO Box 105, A-1014 Vienna, Austria).

2-4 July, **Cerebrovascular Diseases: New Trends in Surgical and Medical Aspects,** Gardonne Riviera (Fondazione Giovanni Lorenzini, Via Monte Napoleone, 23, 20121 Milan, Italy).

7-9 July, **Light Microscopy '81,** London (Royal Microscopical Society, 37/38 St Clements, Oxford, UK).

13-17 July, **Cellular Basis of Idiotypy,** New York (Mrs M. L. Brown, Mount Sinai School of Medicine of The City University of New York, One Gustave L. Levy Place, New York, New York 10029).

13-25 July, **Cosmic Ray Conference,** Paris (CEN Saclay, 17th ICRC, 91191 Gif-sur-Yvette, Cedex, France).

15-17 July, **Thermodynamics and Kinetics of Metallurgical Processes,** Bangalore (G.N.K. Iyengar, Indian Institute of Science, Bangalore, India).

15-18 July, **Improving University Teaching,** Tsukuba (University of Maryland University College, University Boulevard at Adelphi Rd, College Park, Maryland 20742).

16-20 July, **ICRF DNA Tumour Virus Meeting,** Cambridge (ICRF, PO Box 123, Lincoln's Inn Fields, London WC2, UK).

21-30 July, **IASPEI 1981,** London, Ontario (A. E. Back, Dept of Geophysics, University of Western Ontario, London, Canada N6A 5B7).

21-27 June, **Mechanisms of Ion Translocation across Membranes,** Naples (Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720).

27-31 July, **Introductory Course in Food Science and Technology,** Weybridge (Mr P. Moss, The National College of Food Technology, University of Reading, St George's Ave, Weybridge, Surrey, UK).

9-14 August, **Teaching Chemistry in a Diverse World,** Maryland (6th International Conference on Chemical Education, Dept of Chemistry, Box 419, University of Maryland, College Park, Maryland 20742).

17-22 August, **High Pressure in Research and Industry,** Uppsala (Carl-Magnus Backman, 8th AIRAPT & 19th EHPRG Conference, Institute of Physical Chemistry, Box 532 S-751 21 Uppsala, Sweden).

17-22 August, **Cybernautics and Systems,** Mexico City (Prof. J. L. Elohim, (Calle) Antonio Sola no. 43, Col. Condesa, Mexico 11 DF, Mexico).

21-22 August, **Exciton-Contraction Coupling in Muscle,** Banff (Dr G. B. Frank, Dept of Pharmacology, The University of Alberta, Edmonton, Canada T6G 2H7).

23-28 August, **Solar World Forum,** Brighton (The Administrator, UK-ISES, 19 Albemarle St, London W1, UK).

23-29 August, **VII International Biophysics Congress** *and* **III Pan-American Biochemistry Congress,** Mexico City (Secretariat, Cerrada de Popocatepetl No. 51-A, Mexico 13, DF Mexico).

24-29 August, **2nd Asian-Pacific Regional Meeting of I.A.U.,** Bandund (Dr B. Hidayat, Bosscha Observatory, Lembang, Java, Indonesia).

25-28 August, **Strategy in Drug Research,** Noordwijkerhout (IUPAC-IUPHAR Symposium, c/o Merck Sharp & Dohme B.V. Waarderweg 39, PO Box 581, 2003 PC Haarlem, The Netherlands).

28 August-1 September, **Embryonic Development: Genes and Cells,** Basle (IX Congress of ISDB, PO Box 141, CH 4007 Basle, Switzerland).

31 August-2 September, **Prostaglandins and Cancer,** Georgetown (Y. Maddox, Dept of Physiology & Biphysics, Georgetown University Medical Center, Washington DC 2000).

1-4 September, **Olefin Metathesis,** Belfast (Prof. K.J. Irvin, Dept of Chemistry, The Queen's University of Belfast, Belfast, UK).

1-5 September, **Central Nervous System Regeneration,** Catania (Dr B. Harber, Marine Biomedical Institute, 200 University Blvd, Galveston, Texas 77550).

2-4 September, **Environment and Safety,** Wembley (The Organisers, I.E. & S. Conference and Exhibition, Newgate, Sandpit Lane, St Albans, Herts, UK).

2-4 September, **6th Annual Symposium of the Uranium Institute,** London (Mrs C. Wilson, Conference Associates UIS, 34 Stanford Rd, London W8, UK).

7-11 September, **NMR Spectroscopy — Principles and Modern Practice,** Nottingham (Miss L. Hart, Royal Society of Chemistry, 30 Russell Square, London WC1, UK).

8-10 September, **Colloid Science,** Bristol (Miss L. Hart, Royal Society of Chemistry, 30 Russell Square, London WC1, UK).

8-11 September, **Remote Sensing: Down to Earth Management,** Winnipeg (Mr D. Pearson, Registration, Box 1106, Winnipeg, Manitoba, Canada R3C 2X4).

13-18 September, **Drugs and Alcohol,** Jerusalem (Drugs and Alcohol Congress, PO Box 394, Tel-Aviv, Israel).

14-18 September, **Microscopic Aspects of Adhesion and Lubrication,** Paris (Société de Chimie Physique, 10 Rue Vauquelin, 75231 Paris Cedex 05, France).

21-24 September, **Particle Size Analysis,** Loughborough (P.J. Lloyd, Particle Technology Group, Chemical Engineering Dept, University of Technology, Loughborough, Leicester).

22-25 September, **Nuclear Quadrupole Resonance Spectroscopy,** Moscow (6th ISNQRS, Institute of Crystallography, USSR Academy of Sciences, 59 Leninsky Prospect, Moscow 117333, USSR).

22-23 September, **European Working Group on Cystic Fibrosis,** Berne (R. Kraemer, Dept of Paediatrics, Inselspital, CH-30130 Berne, Switzerland).

25 September, **Modern Meat Technology — Microbial Considerations,** Kansas (Dr D.Y.C. Fung, Leland Call Hall, Manhattan, Kansas 66506).

25-27 September, **Light-induced Mutagenesis,** Bath (Dr S.H. Moss, School of Pharmacy and Pharmacology, University of Bath, Bath, Avon, UK).

nature

19 February 1981 19/2/82

Employment, training and opportunity

Economic recessions, painful though they may be, are supposed in the conventional wisdom also to be opportunities for industrial change. Industries whose competitiveness has been undermined close factories, throw workers into the dole queues and perhaps go out of business. In due course, the theory goes, some of the people thus displaced from work become the entrepreneurs who found new kinds of industries; others learn new skills so that when the economy revives they too can join the workforce in a new pattern of industry. Adversity, the textbooks say, can thus be an instrument of rapid change. So does it follow that the wave of recession now sweeping the industrialized West is merely a sign of better things to come? Not if Britain is an example, or if the British government's capacity to make the best of the opportunities now presented is typified by its shabby performance in the debate on industrial training in the House of Commons last week.

Nobody disputes the severity of the recession now afflicting the British economy. Manufacturing output has declined by some 15 per cent since its peak two years ago. Employment in the private sector has declined even faster (which promises increased productivity when the worst is over). Unemployment now exceeds 2.5 million, is most serious among young people and threatens to go on increasing at least until the summer. So is this not a splendid opportunity for making sure that the people out of work are well prepared for the recovery that may yet come? Or at least for giving them incentives to acquire new skills if they are so inclined? Both the government and the Labour opposition seem to agree, if last week's debate is to be taken seriously. Their good intentions are not matched by an imaginative appraisal of what should be done.

The narrow issue that divides them is the future of the network of industrial training boards, first set up in 1964 and financed by means of a levy on the turnover of all companies operating in specified groups of industries. The system is not as onerous as it may seem. Intelligent companies ensure that they recover at least as much in financial support for their internal training operations as they pay by way of levy. Some of the best apprentice schools in Britain are now financed by means of this genteel laundering of money, and the system may also blunt the edge of the longstanding discontent of progressive companies that much of the training they provide eventually benefits less farsighted competitors which pick up skilled men in the job market. The system of training boards has probably also served a useful purpose in industries such as catering, where individual companies tend to be small and where the labour force is mobile. The system, nevertheless, has serious weaknesses. From the start, the training boards have been inclined to impose on the industries they supposedly serve patterns of training whose utility is disputable. Since 1973, when the administration costs of the training boards became a charge on the central government, the boards have become even more bureaucratic than they were before, and extravagant to boot. (At least one director of a training board is paid more than a university vice-chancellor at a British university.)

For several months, it has been clear that this state of affairs could not continue. Even the agencies of the British government sympathetic to the training boards (the Manpower Services Commission, for example) have acknowledged that the system has serious faults. Industrial companies, especially progressive companies, have become increasingly discontented with the arrangements. A few months ago, the Conservative Party's Centre for Policy Studies advocated the simple abolition of the training boards. The Conservative government, as much concerned with the cost of keeping the boards in being (more than £50 million a year) as with the need for training, has resolved to transfer the cost of administration back to the industries which contribute to the boards. The case for this course of action, put forward in the House of Commons last week, is that a transfer of full responsibility for the boards back to the industries concerned would provide a built-in assurance that industrial training is geared to industrial needs.

The argument, so far as it goes, makes sense. It is, however, disingenuous to suppose that there will be no consequential changes. Many industries now saddled, as they see it, with training boards will regard the Employment and Training Bill now working its way through the House of Commons as a licence to do away with their training boards. In many cases, the result would be far from tragic. Many of the complaints recently levelled at the boards are fair. One serious weakness is that there has never been an objective evaluation of the educational value of their training schemes. A more serious weakness is structural. If the best outcome of the present recession of the British economy is the emergence of entirely novel industries, what purpose can be served in perpetuating a pattern of training geared to industries in decline? Although true, none of it excuses the narrowness of the British government's case for change.

Everybody accepts that this point in the current business cycle calls for new initiatives in training. The British government's argument that the responsibility should as a matter of principle fall squarely on industry is, however, profoundly flawed. Progressive industrial companies have always acknowledged that their own self-interest requires some investment in the skills of their own workpeople, and will continue to do so. Most of them acknowledge that the cost is worthwhile, even if some of their skilled people eventually work for competing companies. But the same industrial companies have no interest in providing their workpeople with skills suited to entirely novel industries. The result is that people working in, say, traditional mechanical engineering are unlikely to be given the chance, within their own organizations, of becoming skilled in electronic engineering or computing. And what of those who are out of work, and thus outside the scope of the industrial training boards? In Britain, in the past few years, the Manpower Services Commission has done some useful work in keeping young people off the streets and providing them with modest skills, but none of this amounts to the creation of the body of skill on which an entirely novel pattern of industry might be built some years from now.

One paradox is that a government devoted, so far as can be told from its public declarations, to individual initiative should so overlook what is now the obvious need — that people by their individual decisions should more easily be able to acquire the skills they consider would help them to make their way in the brave new world ahead. Another paradox is that there is no shortage of facilities for the provision of training, and, by all accounts, no shortage of people seeking skills. Both the universities and the polytechnics are at present short of students and, faced with declining support for conventional students' education, more eager than ever to provide vocational training of various kinds. So why does not the potential demand for advanced training more fully meet the potential supply? At least one part of the explanation is the cost. Those in jobs must pay the full cost of a training course on which they embark out of their taxed income. Those out of jobs are usually unable to raise the

0028-0836/81/080619-02$01.00

fees at all. The implications are plain. If the British government means what it says about the importance of training for the future health of British industry, it must make the cost of training fees paid out of a taxpayer's pocket tax-deductible. And it must use some of the £50 million a year it will save on the industrial training boards to help those out of work to acquire new skills.

In the past two years, the British government has been outwardly resolute in sticking to the monetarist guns that have made the present recession the painful phenomenon it is. The government has been alternatively praised for courage and pilloried for pigheadedness. In reality, it has followed an intermediate course — nationalized industries such as British Leyland and British Steel have been saved from collapse by taxpayers' money. Universities and other institutions of higher education have been exposed more directly to the cold winds of economic change, and like many parts of British manufacturing industry, are running well below capacity. They too will emerge transformed from the present slump. One possible and desirable outcome would be that at least some of them should be more directly involved with continuing education. If the government means what it has been saying these past several months about the importance of retraining and about its respect (or at least lack of enmity) for higher education, and because the falling inflation rate suggests that the time is fast approaching when reflationary policies will again be necessary, it will be absurd if this opportunity is let slip.

How to make dreams untrue

Although the hunt for the philosopher's stone has long since been abandoned, the hunt for its antidote, the material that turns everything into lead, continues apace, most successfully in the dealings between the British government and the nationalized industries for which it is responsible. This is the chief lesson to be drawn from the report of the House of Commons Select Committee on Energy (see page 621) on the new British nuclear energy programme. The committee plainly found its inquiry into the rights an wrongs of the proposal and infuriating experience. The British government chose not to pass on to the committee the advice provided by the Central Policy Review Staffs on the wisdom of building two reactors of the advanced gas-cooled type, thus confirming the widespread understanding that the advice was hostile to the project — and the belief that when British governments extol the benefits of open government, they are thinking of governments other than themselves. The Department of Energy, ostensibly in charge of the substantial commitment of resources entailed by the nuclear programme, is convicted of incompetence by its evidence in the House of Commons. The Central Electricity Generating Board, the nationalized electricity utility which is the constructor and operator of nuclear plants south of the Scottish border, is accused not of excusable incompetence but of unforgivable deviousness. When all the dust has settled, the chief casualty of the committee's forthright report will be the British nuclear programme, fanned to life only in 1979 from the embers of a decade's neglect.

Even after due allowance is made for the way in which select committees these days have to shout to be heard above the reports that flood from their competitors, the language used by the Select Committee on Energy will surprise many of those at whom the complaints are levelled. For the committee's chairman and deputy chairman are both members of the now defunct Select Committee on Science and Technology which, over the years, has been one of the most consistent cheerleaders for nuclear power in Britain. What has gone wrong, the luckless targets of the report will be asking? The answer is that they have compromised a good cause by arrogance, fecklessness (which is a synonym for conscious laziness) and a failure to understand what they are for. In the short run, only they may suffer. Further ahead, taxpayers (and electricity consumers) will also be the losers.

The problem of nuclear power in Britain is by now a problem in national mythology. Ever since the Tizard mission to Washington in 1941, the popular belief in Britain has been that nuclear fission was invented ("discovered" is too passive a word) in the land of Newton. The Ottawa Agreement in 1944 merely codified what most people knew, for which reason it followed as night does day that there should be a truly British nuclear explosion (in 1948) and, now, a nuclear deterrent (called Trident) whose design may be American but whose construction is beyond question British. The commercial exploitation of nuclear power in the past quarter of a century has been a cross between a fairy story and a nightmare. From time to time, radical innovations of technology have been announced, have been widely advertised to an unsuspecting press and then have been allowed to run into the sand because nobody cared about them in the sense of knowing that not merely his reputation but his pension (index-linked) would depend on the realization of his promises to the public. To everybody's surprise, however, the nuclear reactors built in the first heady days of the nuclear enterprise (with natural uranium as fuel and graphite as a moderator) are mostly still working well and helping to reduce the price paid by electricity consumers.

The select committee's forgivable irritation has plainly led it to complain most vociferously at those who sought unsuccessfully to mislead it — the Central Electricity Generating Board is the worst offender. The issue, the committee says, is whether the board can possibly be right in pretending that the nuclear power stations it plans to order in the coming decade will be economical devices for generating electricity when there is so much evidence that the board's assumptions are either untested or untestable. All kinds of questions arise. Can the board be right in guessing that the cost of producing coal (the obvious alternative to uranium) will increase by between 2.5 and 3 per cent a year in the foreseeable future? Can it be believed in quoting capital costs for building the simplest structures when it has so mismanaged the oil-fired power station still painfully half complete on the south bank of the Thames Estuary? Is it, indeed, possible to consider seriously the figures produced by the board and others, purporting to quote future electricity generating costs, if none of the interested parties turns a hair when it turns out that the estimates are wrong, and that the consumers of electricity must pay the difference? For all its rhetoric, the committee would have had a more telling case if it had been able to specify the premises on which utilities (such as the Central Electricity Generating Board) might set out to calculate the future cost of electricity. The truth is, of course, that in the real world there are none that carry conviction. What happens if there is, in the end, a war in the Middle East? It would be too late, then, to call the committee back into session.

With luck, such a drastic course will not be necessary. The committee has done what it needs to do for the decade ahead. It has thrown a spanner in works that were not working. The recommendation that the United Kingdom Atomic Energy Authority should be the chief agent of technological change in the nuclear industry, concerned principally with fast reactors and with fusion, is to be welcomed. The plea that the Department of Energy should somehow equip itself to understand the policies with which it pretends to grapple is right, but will quickly be forgotten with the connivance of the department. The crux of the issue, with which even this loudspoken committee has not come to grips, is that there is no mechanism in the present arrangement for the sponsorship of nuclear or any other kind of power that will allow that the consumer of the electricity should be king. The present British government has widely advertised its wish to reduce the public sector's borrowing requirement by inducing private investors to share some of the risk. Is it not time that it sold the Central Electricity Generating Board to somebody with a long purse — or better still, a conscience? The committee's irritation has led it correctly to pillory the board. It should go on to define criteria by which all such boards should measure their duty.

More secrecy on cryptography research

US academics lean towards self-restraint

Washington

Mathematicians and computer scientists in the United States will soon be asked voluntarily to submit papers on cryptography and related research to the National Security Agency before publication, to see if the agency feels they contain anything that should be kept secret.

This precedent-setting system of self-censorship is being proposed by a study group set up last year by the American Council on Education to look at the growing conflict between national security and academic freedom in cryptography research.

The proposal is something of a compromise between the Defense Department's demands for strict legislation to control the publication of research results with potential security implications, and critics who argue that there should be no restriction on the publication of non-classified research.

However, several scientists warned last week that such voluntary self-regulation could lead eventually to demands for a similar approach to research ranging from lasers to integrated circuits.

Tensions between the Defense Department's security agency and sectors of the research community have grown steadily over the past few years. They stem partly from a desire by the agency to limit the spread of knowledge about virtually unbreakable codes — and the insistence of mathematicians that since many difficult mathematical problems can provide the basis for such codes, any restrictions would have a "chilling" effect on research.

Vice-Admiral Bobby Inman, director of the National Security Agency, has argued fiercely that the free publication of research results could inhibit the agency's data-gathering capabilities. Scientists reply that the codes also have important civilian applications — such as the protection of computer data — that justify their wide dissemination.

Last year these tensions rose to the surface when a computer scientist at Massachusetts Institute of Technology, Dr Leonard Adelman, found a grant application to the National Science Foundation had been passed to the National Security Agency. The agency subsequently offered to support part of his research — but on terms which would have given it the right to determine how much should be published.

The incident caused considerable embarrassment to the National Science Foundation — which protested that it had been seeking the views of the security agency on cryptography research applications for several years. It also pointed out that potential conflicts were being studied by the study group of the American Council on Education, set up at the suggestion of the National Security Agency to discuss ways of controlling the distribution of research results acceptable to the scientific community.

After a year's study, the group agreed at a meeting in Washington last week to propose a system leaving responsibility in the hands of scientists and journal editors by setting up a voluntary review system by the security agency.

According to the group's proposals, soon to be circulated in the scientific community, either a scientist or an editor could submit a paper to the agency for comments on whether it contains information considered to be a threat to national security.

If the agency had no objection, the scientist would be free to publish. If it did object, then the scientist could decide not to publish, proceed with publication against National Security Agency advice — or refer the matter to an independent, five-person committee. This would have two individuals named by the security agency, and three picked by the President's science adviser from a list submitted by the National Academy of Sciences.

In principle, officials of the American Council on Education hope that voluntary self-regulation — which would be introduced for a trial period — would avoid the difficulties of new legislation (which could run into constitutional problems over freedom of expression) while meeting the security agency's main concerns.

In practice, getting the system to work will not be easy. The first step will be for the security agency to prepare a guide to the type of research projects it would expect to evaluate. If the list is too broad, agency officials admit they could end up stifling research unnecessarily; yet if it is too narrow, they fear both that they might tip off others about their principal interests, and miss potentially valuable research findings.

There is also likely to be considerable resistance from the scientific community. Only one of the study group's nine members voted against the proposal for self-censorship; this was Dr George Davida of Georgia Institute of Technology, who found a patent application intercepted by the National Security Agency three years ago, and subsequently received a letter threatening consequences if he discussed his research with his colleagues.

Several academics, however, are worried that self-regulation would create a new category of secret research, pointing out that classified research is now banned on many campuses following the anti-war demonstrations of the 1960s.

The study group's proposals are therefore likely to generate considerable heat. But the political tide is now running in its favour and those who protest at the encroachment of security agencies on individual liberties have fewer friends in Congress than in the past. Vice-Admiral Inman has been nominated as deputy director of the Central Intelligence Agency — and remains committed to the desirability of strong government controls over potentially sensitive research.

David Dickson

Committee douches nuclear energy

The British government's 1979 statement on nuclear power, like its predecessors, is a muddle. This is the opinion of the House of Commons Select Committee on Energy, published this week. The committee asks that decisions to build nuclear plants in the 1980s and 1990s should be decided on their merits and not as part of a planned programme.

On economic grounds, the committee is sceptical about the government's programme to build 15 GW of new nuclear plant by 1992. The Central Electricity Generating Board (CEGB) comes in for particularly sharp criticism. The report cites several instances where the board's evidence on costs was misleading. It criticizes the board for basing future costs on early Magnox plants without acknowledging the effects of subsequent inflation on future capital investment, and for comparing the costs of electricity generated by different types of plant by using "highly uncertain variables" such as the average load factor of plant and future fuel and fuel cycle costs.

Most damning is the complaint that the CEGB presented international cost comparisons suggesting that a pressurized water reactor (PWR) would cost 34 per cent more to build in Britain than elsewhere. The committee says that the generating board's estimate of PWR costs are "too perfunctory" and that it is too tolerant of inefficiencies in the British construction industry. Planning permission for the first PWR plant is still to be sought. Subsequent plants will be either PWR or AGR (advanced gas-cooled reactor) depending on cost and performance.

The committee also suggests that the size of the British programme could be cut if CEGB and the South of Scotland Electricity Board reduced their planning margins for the excess capacity needed for plant failure in particularly severe winters. These have crept up to 28 and 73 per cent respectively from about 17 per cent in the

0028-0836/81/080621-02$01.00

fifteen years ago.

The electricity generating boards are not the only organizations to be criticized in the report, however. Criticisms of the Nuclear Installations Inspectorate, the independent body responsible for nuclear safety, made in Parliament last year, is reiterated. The inspectorate needs to put more effort into assessing the PWR, a design new to Britain, says the committee. Two aspects of PWR design, the integrity of the pressure vessel and problems of two phase flow in the water coolant, call for highly specialized inspectors which the inspectorate lacks. The committee recommends that the inspectorate takes on an ultrasonics expert for testing pressure vessels and that the government remedies, by means of legislation if necessary, the inspectorate's difficulty in attracting suitably qualified staff because of uncompetitive salaries.

The UK Atomic Energy Authority (UKAEA) is criticized for past bad advice to government. The committee recommends that its role as adviser to government on nuclear policy be given to the Chief Scientist's office at the Department of Energy. The UKAEA, it says, should confine itself to research on future nuclear options such as the fast breeder reactor and fusion.

One example, in the committee's opinion, was the recommendation not to consider the Canadian CANDU reactor as a possible option. The committee clearly believes that the CANDU reactor offers certain advantages over both the AGR and the PWR but realizes that the government is too committed to its present plan for an about-turn now. Instead, it asks for a study of CANDU before a final commitment to the PWR is made. **Judy Redfearn**

European Community

Research compared

Brussels

The European Community seems to have taken fright at the gap between Europe and the United States and Japan in spending on research and development. In his first speech to the European Parliament last week, the president of the European Commission, Gaston Thorn, said that in 1981 the Commission will give priority to research and development that will improve the Community's competitiveness with Japan and the United States. The latest assessment of research spending shows that in Europe as a whole, research and development expenditure accounts for 1.9 per cent of Gross Domestic Product, compared with 2.3 per cent in the United States and 2.0 per cent in Japan. If defence research is excluded, the figures are 1.7 per cent, 1.7 per cent and 2.0 per cent, indicating Japan's present dominance by this yardstick.

What, asked Vincent Ansquer, a French member of the European Parliament, can the Commission do? The answer seems to be that the Commission will hope to increase its own research spending in the next few years from 1.6 per cent to 2.0 per cent of the combined research and development budgets of the member states.

Constitutionally, the Commission is less able to influence national expenditure directly but is working on a series of studies which are to be summarized as guidelines for a common research and development policy due to be presented at the Council of Ministers in June this year.

Community research and development expenditure is assessed each year in a report by the Scientific and Technical Research Committee (CREST). Comparisons between member states have been quite influential in the past; for example, they prompted the French government to make substantial increases in its research budget in both 1979 and 1981. The latest report, still being finalized, will say that between 1979 and 1980, member governments' expenditure on research grew in real terms by at most 0.4 per cent.

During the whole of the 1970s, it is now clear, growth rates were highest in West Germany, Ireland and the Netherlands, with France, Italy and the United Kingdom below average. In 1980, Italy emerged at the top of the growth table, with an increase of 20 per cent in real terms of its research and development budget.

The CREST report uses information about the objectives of government expenditure in 1970–79 to infer changing priorities. Throughout the Community, government support for "the general promotion of knowledge", still the largest item, is declining. In stark contrast to government declarations, however, European governments seem to have allowed research contributing to industrial productivity and technological development to fall proportionally.

Increased proportions of national budgets have been spent on the exploration of the Earth and the atmosphere, the planning of the human environment and the protection and improvement of human health. The widely acknowledged need to reduce European dependence on imported energy seems not to have received its due — over the 1970s, real expenditure increased by a mere 0.4 per cent and actually decreased by 1 per cent immediately after the increase of oil prices in 1973–74. Energy research, nevertheless, has an important role in Italy and Germany, while in Denmark and the Netherlands most government expenditure continues to be "on the general promotion of knowledge".

The influence of defence research expenditure on the pattern of research in individual countries is catalogued in the CREST report. The disparity between the United Kingdom and its partners stands out. Thus British expenditure on defence research grew from 41 per cent of the government's budget in 1970 to 53.3 per cent in 1979. Only France comes anywhere near this proportion, with 35 per cent of the government's budget going to defence research in 1979. In the same year, Germany spent 11.7 per cent on defence research and the other six member countries little or nothing. The report does point out, however, that defence research brings industrial spin-off.

The outlines of the committee's picture remain clear enough. Research spending by Community governments is now increasing after the trough in 1978 but still falls short of American expenditure. (The Japanese statistics are incomplete.) In money terms, research and development expenditure in the United States in 1979 was about 1.3 times greater than the corresponding figure for the nine members of the Community. The good news, though, is that in 1980 (according to forecasts of US federal agencies) American government expenditure may have declined while that in Europe remained constant. **Jasper Becker**

Toxic chemicals

UK regulations

The long awaited and potentially contentious draft regulations by which the British government will require the notification of new chemicals were published by the Health and Safety Commission on Wednesday (18 February). The provisional timetable for discussion allows until July for comment, in which case the regulation could become law before the deadline of 18 September laid down by the European Commission. But Brussels, in this as in other matters, is running late and may not be in a position to make its 1979 directive binding on all member states until next year.

The draft regulations are substantially more stringent than the proposals described in the commission's discussion document published in 1977. The principle that they should apply only to new substances remains, but outline notification will now be required for substances manufactured or used in quantites of less than 1 tonne. Chemical manufacturers will now also be required to provide information about biodegradability, thus meeting one of the criticisms of the earlier proposals by environmentalists. One measure of the increased stringency of the draft regulations is that the estimated cost of testing a new chemical is now given as £45,000.

The proposed regulations for the United Kingdom are more stringent than those required by the European Community in that they apply to chemical intermediates as well as to products supplied to others. This is one of the points on which British chemical manufacturers are likely to concentrate in the coming months, but there will also be complaints that the proposed regulations allow the Health and Safety Executive to pass on to other European authorities information about

0028-0836/81/080622-02$01.00

chemicals newly introduced by British manufacturers.

Full notification (from which chemicals supplied to research laboratories will ordinarily be exempt) is elaborate. Manufacturers will be required by means of an acute toxicity test to classify new substances as toxic (LD_{50} less than 200 mg per kg), "very toxic" (LD_{50} less than 25 per kg) or neither, using oral doses in laboratory rats. Ordinarily, the chemical identity of the material will have to be defined, as will the purity of the material and "the nature and the concentration of the main impurities", which may be harder.

For most manufacturers, biological tests will be the most time-consuming. Apart from acute toxicity data (with oral doses in rats, subcutaneous doses in rats or rabbits and inhalation doses in rats), notification requires a 28-day sub-acute toxicity test, acute toxicity tests with fish and daphnia, tests for skin and eye irritation and of skin sensitization and a mutagenicity test. The regulations empower the Health and Safety Executive to require, on the basis of a preliminary examination of the first set of data (which has to be completed within 45 days) further information including the results of tests for teratogenicity and carcinogenicity.

Exceptions to the strict requirement of notification include the supply of chemicals required for further development and materials supplied in quantities amounting to less than 1 tonne (where chemical identity will usually be sufficient). One possibly attractive let-out is the provision that users of such small quantities may be able to meet the requirements by submitted a sample in a container labelled "CAUTION — substance not yet fully tested".

Industrial innovation

States chip in

Washington

Support by the states for technological innovation is becoming fashionable. Echoing a theme launched in Congress and by the federal government three years ago, several states are now beginning to put political force behind legislation intended to stimulate technological innovation in local industry. Many realize that their economic future may depend on their ability to attract and keep high technology companies in competition with rival bidders. Measures are therefore being introduced at the state level comparable with the federal initiatives proposed by President Carter two years ago for stimulating investment in high technology.

Often the new measures include the provision of substantial research funds for state universities to promote research in areas of particular commercial value such as the technology of microcomputers. Many states are also trying to draw their universities closer to the research and training needs of the private sector. Last week, for example, the state assembly of New York passed a package of four bills to promote the growth of high technology industry in the state and to forge closer links between the state government, private industry and local universities.

One of the New York bills would revivify the state's Science and Technology Foundation, originally established by Governor Nelson Rockefeller in 1973 to improve the quality of university research in the state. The new bill, likely to be accepted by the state senate and passed into law, would require the foundation, now virtually defunct, to help small companies obtain federal research grants such as those offered by the National Science Foundation. The bill also provides new incentives for colleges and universities to carry out potentially commercial research, and two members with management experience in high technology companies would be added to the foundation's board. A second bill, concerned specifically with research workers at the State University of New York, would increase the proportion of profits from the successful commercial development of research results that might be kept by a faculty member.

The growing interest of all states in the development of industrial policy has prompted a synposium organized this Saturday (21 February) in Washington by the governor of California, Edmund G. (Gerry) Brown immediately before the annual meeting of the National Governors' Association. Mr Brown has recently proposed a $22 million "reindustrialization" programme for California, claiming the need to forge a new partnership between government and industry. This demand was the basis of his unsuccessful bid for the Democratic presidential nomination last year — and is expected to emerge again in the 1984 election.

Central to Mr Brown's programme is a proposal to establish a Microelectronics Innovation and Computer Research Operation. This would be based at the Berkeley campus of the University of California and would focus on basic research in microcomputers considered too costly to be carried out by individual companies. The state would allocate $7.6 million to the project, $2.6 million to be spent by the university and the rest to be matched by grants from private companies, most of which have welcomed the proposal.

Similar plans are being worked up by other states. For example, the Minnesota state legislature is considering a request for substantially increased funds from the Microelectronics and Information Sciences Center at the state university. In North Carolina, whose governor, James D. Hunt, has promised to make the state the centre of the East Coast microelectronics industry, the government agreed last summer to allocate $1.8 million to a microelectronics centre in Research Triangle Park which has already tempted General Electric to commit $50 million to an integrated circuit plant nearby.

The Californian initiative is in part defensive. Because of high real estate prices and shortages of technical staff, many states are now sending delegations to Silicon Valley to persuade established semiconductor companies of the virtues of relocating elsewhere. Ohio has already been offering attractive inducements to companies prepared to move their operations, and while Mr Brown will be talking in Washington on Saturday, a delegation from New Orleans will be in California describing the advantages of the Deep South. **David Dickson**

Telecommunications

Monopoly intact

Halfway through the committee stage in the House of Commons, the British government has agreed to amend its telecommunications bill in ways that will please British Telecom but are unlikely to reassure telecommunications equipment manufacturers. The amendment will give British Telecom, the new telecommunications authority, the right to be consulted during licensing of equipment for attachment to the public telephone network. Manufacturers, already dismayed at the modest erosion of the public monopoly promised by the original version of the bill, now fear that their freedom to develop new communications technology will be further constrained.

The bill, introduced by Sir Keith Joseph, Secretary of State for Industry, last year, is the British government's watered-down equivalent of the steps taken by the US Federal Communications Commission to stimulate innovation by closer definition of the public monopoly in telephone systems. (Similar moves are under way in West Germany.) The British bill, which would split the mail-handling and telecommunications parts of the Post Office into two separate corporations, would leave British Telecom with total control of the telephone network but shade its monopoly on the supply of equipment.

The procedures in the original version of the bill for ensuring that attachments do not interfere with the operation of the network would take control from one bureaucracy, the Post Office, and give it to another, the Department of Industry. One of the amendments has it that standards will be set by the British Standards Institution in conjunction with British Telecom and the department; licences to manufacture equipment are now to be issued by the department after taking British Telecom's advice. Manufacturers complain that standards will take a long time to set and that British Telecom, potentially a manufacturer in its own right, will retain too much control. These issues are likely to be debated further before the

0028-0836/81/080623-02$01.00

committee stage ends, probably before the Easter recess.

Under the new arrangements, British Telecom will be able to manufacture and sell attachments only through independent subsidiary companies. Financing such ventures, however, will be difficult within the government's constraints on public borrowing. This is one reason why the government has amended the bill to allow British Telecom access to private finance. It may allow the corporation to work outside the public sector borrowing requirement, an issue not finally settled by the Treasury. One difficulty is that, whatever the source of capital, the government is still ultimately the guarantor.

In the next few months, equipment manufacturers will be keenly waiting to see how the removal of the monopoly is to be staged. The plan is to have a transitional period of three years during which the Department of Industry would control the market, giving British manufacturers time to plan for free competition. Much will depend on the Secretary of State for Industry, who is given wide enabling powers by the bill. Another open question is the arrangements for value added services. The bill gives the Secretary of State power to instruct British Telecom to allow private companies to lease part of the public network for the provision of services to third parties. What Sir Keith finally decides may depend on the outcome of a study by Professor Michael Beesley of the London Business School. **Judy Redfearn**

Carcinogen criteria

US retracts

Washington

Pressed by a recent Supreme Court ruling, US occupational safety officials have agreed to relax their previous opposition to including risk assessment calculations in decisions about reducing exposure to potential carcinogens. Hitherto, officials of the Department of Labor's Occupational Safety and Health Administration (OSHA) have argued that the scientific uncertainty of risk calculations — and the lack of an explicit requirement in the agency's authorizing legislation — meant that they should not be used as the basis of policy decisions. Instead, they have maintained that that once a substance has been demonstrated as a potential hazard, exposure should be lowered to the "lowest feasible level".

OSHA is being required to shift its position as a consequence of a decision by the Supreme Court last summer to strike down regulations which the agency had proposed for limiting industrial exposure to benzene (*Nature* **286**, 97; 1980). The court ruled that OSHA had failed to demonstrate that reducing benzene exposure limits from ten to one parts per million would reduce the level of "significant risk". Following this ruling, OSHA has now announced that it is revising its generic carcinogen policy, introduced last year as a mechanism for regulating any chemical suspected of causing cancer.

Under the terms of this policy, any chemical which meets one of a number of criteria — for example, which is shown to cause cancer in one experiment with laboratory animals, and also produces a positive result in a short-term test — is automatically labelled a carcinogen.

Originally, this would have been sufficient to invoke automatically the requirement that exposure be reduced to the lowest feasible level. Now OSHA officials have agreed to include consideration of whether the chemical poses a significant risk — using a variety of data to make this judgement, including court interpretations of previous rulings, OSHA's previous experience in regulating toxic substances and "prudent occupational health policy".

According to an announcement made in the *Federal Register*, three aspects of the benzene decision will be incorporated in the cancer policy. First, the significance of existing risk must be estimated before issuing a carcinogen standard; second, the exposure level must be set at the lowest feasible level which is "reasonably necessary or appropriate to eliminate significant risk"; and third, that OSHA must consider "all relevant evidence" in making such determinations.

This change in policy represents a significant shift from OSHA's previous position. The agency's former head, Professor Eula Bingham, has described risk assessment calculations as "abhorrent" to public health administrators; and the agency has strongly resisted pressure from government economists to push occupational regulations into a neatly quantifiable mould.

At the same time, the shift does not go as far as many in industry would like. Their demand is for full-scale cost-benefit analysis of all occupational health and safety regulations, based on the argument that federal controls have become a major economic burden. Although the Supreme Court, in striking down the benzene regulation, did not make the widely expected pronouncement on whether cost–benefit was required to demonstrate that a new regulation was "reasonably necessary", it is expected to do so in ruling on another case which has been brought against OSHA on cotton dust standards.

In any case, agency officials expect that a demand for full-scale cost-benefit analyses of future regulations — with the requirement that the least expensive option be adopted — will be one of the first ways in which the new Administration will try to meet its election promise of reducing the force of government regulations.

The labour movement, which has consistently argued that cost–benefit analysis is little more than a smokescreen designed to cover the relaxation of safety and health controls, is preparing for a fierce and lengthy battle. It has already complained about one of Mr Reagan's first anti-regulation acts, withdrawing new regulations published by OSHA in the last days of the Carter presidency which would have required the labelling of all hazardous chemicals used in the workplace.

David Dickson

Greenhouse effect

Act now, not later

Stockholm

The theme that the time has come for policy-makers to take account of carbon dioxide when drawing up energy policies ran through an Earthscan meeting on carbon dioxide, climate and, energy last week. But speakers' conviction that action should be taken now was matched by their caution in predicting exactly what would happen if carbon dioxide emissions continue to increase.

The fundamentals are broadly agreed. The pre-Industrial Revolution atmospheric concentration of carbon dioxide was about 290 parts per million, and the burning of fossil fuel has been the largest single factor contributing to concentrations which are expected to double by the middle of the next century, assuming a 2 per cent annual growth rate in the use of fossil fuels.

The reality of the greenhouse effect was also common ground between the speakers. Predictions about specific climatic changes in specific parts of the globe were, however, more equivocal.

Current models, according to Professor Bert Bolin of the University of Stockholm, are inadequate but "they are all we have". The models are especially inadequate in dealing with the role of clouds and the interaction of the atmosphere and the oceans. Dr Tom Wigley of the University of East Anglia pointed out the difficulties of distinguishing the signal from the noise: knowing when variations in regional climates stem from a particular factor such as carbon dioxide and when they are simply part of the continual natural variation.

Professor S.K. Sinha from the Indian Agricultural Research Institute in New Delhi was the only speaker daring to be at all optimistic, and even his belief that agriculture could adapt to climate changes was conditional on fruitful research being done on water management, the identification of new genotypes more tolerant of temperatures 3–4°C greater than at present and on higher crop yields with a smaller input of fossil fuels.

The most eloquent plea for energy policies to take account of carbon dioxide came from Gus Speth, chairman of President Carter's Council on Environmental Quality. In the last days of the Carter presidency, the council urged that "full consideration" should be given to carbon dioxide in the development of

0028-0836/81/080624-02$01.00

United States and global energy policy.

One of the first priorities in a preventive strategy would be to decide "what level of atmospheric carbon dioxide should be considered a prudent upper bound". Should we allow an atmospheric build-up of, say, 50 or 100 per cent over pre-industrial levels? The upper bound would carry with it implications for both developed and developing countries and would raise questions about the sharing of the fossil fuels whose use would be allowed.

Speth was not particularly optimistic about the chances of getting these things done. It is very hard to provoke an international response to an intangible problem whose consequences are not yet even predictable. Dr Thomas B. Johansson of the University of Lund could offer a little encouragement in that the sorts of energy policies desirable from a carbon dioxide point of view were also becoming increasingly necessary, in Sweden at least, for economic reasons. **Wendy Barnaby**

Plasma research

German setback

The Max-Planck-Institute for Plasma Physics at Garching has been forced to make a major reappraisal of its future. Because of government financial restraints (*Nature* 5 February), the institute's next major project, the "Zephyr" tokamak, has been cancelled. The institute, one of the three leading fusion laboratories in Europe, is now planning that its next project should be financed by a redistribution of resources within its total budget of about DM100 million a year.

Zephyr had been planned to leapfrog the joint European tokamak machine called JET, now being built at Culham in the United Kingdom, and would have experimented with ignited plasma. The hope now is that a redistribution of the budget will yield between DM20 million and DM100 million over the next seven years to build a less ambitious machine.

The reassessment at the institute will be carried out under new management. Last week, Professor Klaus Pinkau was appointed director of the laboratory. Although not a plasma physicist but a cosmic ray physicist, he has considerable experience of international collaboration, the politics of big science and the management of scientific institutions. He has been the director of the Max-Planck-Institute for Extraterrestrial Physics, next door at Garching, and is chairman of a committee reporting to the federal government on the merits of ten big science projects which, curiously, did not include Zephyr.

Pinkau said last week that it was dangerous to make scientific institutions too dependent on "annual changes" in the financial position of governments but, also, that budgets should not grow too fast. Certainly this year's changes at the Institute for Plasma Physics will give him pause: not only has Zephyr been deleted but the proposed budget for 1981 has been cut by 15 per cent. This trimming of sails may give the laboratory a sense of realism, persuading it that it cannot alone compete with JET, but a tough internal struggle seems inevitable between the advocates of an upgraded stellarator and a mirror machine. The edge might be taken off this battle if the new project were adopted by Euratom as a "preferred project", in which case between 10 and 20 per cent of the cost could come from Brussels. The advocates of an improved stellarator point to their success last year when the existing machine at Garching, Wendelstein VIIA, was used to show that a stellarator plasma could be held stable in conditions only previously obtained in tokamak machines. The same series of experiments created conditions of plasma density and confinement time more stringent than those reached by tokamaks of similar size, apparently putting stellarators back in business.

Tokamaks are in fact beginning to lose favour because of the various difficulties (fatigue and maintenance, for example) expected to arise in power reactors. Diversification is therefore considered prudent, whence current interest in stellarators and mirror machines. The Lawrence Livermore Laboratory in California is in fact building a mirror machine (the Mirror Fusion Test Facility) in which a large long solenoid is plugged at the ends with magnetic quadrupole mirrors. Garching cannot hope to compete with Livermore in money terms but, some argue, could attack the principles of such a device. This, broadly, is the second proposal being considered at Garching. A decision is expected in the middle of the year. **Robert Walgate**

Tokamaks and stellarators

In both tokamak and stellarator, plasma is confined to a torus by two superposed magnetic fields. One field runs around the torus, along its long circumference; the other winds round and round the small circumference of the plasma ring.

The first field is created in both tokamak and stellarator by a coil wrapped round the small circumference of the torus. The second is created differently in tokamak and stellarator.

In the tokamak, it is the result of a current carried in the plasma itself; in the stellarator it is the result of a component of current in the outer coils along the major circumference of the torus. As the plasma ring is electrically isolated, the tokamak must create the longitudinal current in the plasma by a transformer effect, and so must be pulsed.

The stellarator, on the other hand, can in principle be run statically. Early tokamaks — which were invented in the Soviet Union — were successful principally because of their large "aspect ratio" (major torus radius over minor), it is now believed, rather than because of any intrinsic merit of the tokamak design; and the need for pulsed operation is seen as a disadvantage in the construction of a true reactor, where the variation of thermal, neutron, and magnetic stresses could increase material fatigue.

Princeton perplexities

There are slippery hands and red faces at the Princeton Plasma Physics Laboratory, where a gaggle of lawyers is trying to decide who was responsible for dropping a 350-ton generator component during the construction of the Tokamak Fusion Test Reactor (TFTR).

The accident happened in December, when the outside stator of the vertical axis generator was being lowered into position. A crane bearing broke and the stator fell 15 feet, damaging both itself and the central rotor.

The incident is not expected to have a significant impact on the construction schedule for the TFTR, which will be used to achieve energy breakeven for the first time and to investigate the engineering features of large fusion systems. A second generator, already in place, will be able to supply sufficient energy for the test reactor, at least in the early phases of operation. However, a detailed study will now be necessary to determine whether the generator can be repaired — or whether a replacement is needed, which could cost up to $2 million, and take some time to deliver.

Present construction schedules anticipate that the TFTR will come into full operation in July or August 1982. This is about seven months later than the original completion date of December 1981, due largely to delivery delays on some of the major components — in particular the toroidal and ohmic field coils used to contain the plasma — which have presented more technical difficulties than expected.

Officials at Princeton say that they do not foresee any insuperable problems, as most of the technology is "state of the art". However, the delays will inevitably add to the construction costs, which are expected to exceed the predicted $284 million by about 10 per cent.

More than five subcontractors may be involved in the heated debate over the responsibility for December's accident. A report is expected shortly from the Department of Energy, which is expected to identify errors of judgement responsible for the crane overload. However, with large insurance sums at stake, any such conclusion is likely to be contested — and will almost inevitably end up in the courts.

0028-0836/81/080625-01$01.00

CORRESPONDENCE

Wrong rodents

SIR — The discovery that several cell lines claimed to be derived from patients with Hodgkin's disease in fact come from an owl monkey [1] emphasizes the vital importance of quality control in biological research.

Several cases of genetic contamination of inbred strains of laboratory mice and rats have come to light recently, and although these are not associated with any known scientific fraud, they have led in some cases to a considerable disruption of research projects. In the United States, for example, one large commercial breeder appears to have been selling "inbred" Lewis (LEW) rats which are not histocompatible, leading to a flood of complaints from research workers and inevitably casting doubt on the validity of some published work. In Japan, a survey of over 100 colonies of inbred mice [2] has shown that about 10 per cent had been genetically contaminated either recently (with continued genetic segregation at some loci) or in the more distant past. Indeed many of the major sublines of common inbred strains such as C3H and C57BL have arisen as a result of genetic contamination in the past [3].

In the United Kingdom a voluntary genetic monitoring scheme for commercial breeders was started about five years ago by the Medical Research Council Laboratory Animals Centre and several cases of genetic contamination have been discovered since that time [4]. Both of the main UK suppliers of inbred mice and rats took these results so seriously that they have established their own in-house genetic monitoring programmes run under the direction of consultant geneticists. The International Committee on Laboratory Animal Science (ICLAS) is preparing a handbook on methods of genetic monitoring, and is considering the establishment of some international reference centres for genetic monitoring. In the meantime research workers are strongly urged not to take the authenticity of the strains that they use entirely on trust. If they purchase animals from commercial breeders they should demand to know what steps the breeder is taking to monitor the stock. Those who breed their own strains should remember that accidental contamination can occur at any time, and they should check their stock using skin grafting or some other suitable method (such as the study of biochemical [5] or immunological markers [6] or morphological features such as mandible shape [7]) as a matter of routine.

MICHAEL F.W. FESTING

MRC Laboratory Animals Centre,
Carshalton, Surrey, UK

1. Dickson, D. *Nature* **289**, 227 (1981).
2. Nomura, T. *et al. Proc. 7th ICLAS Symposium, Utrecht* (Fischer, Stuttgart, in the press).
3. Festing, M.F.W. *Inbred Strains in Biomedical Research* (Macmillan, London, 1979).
4. Festing, M.F.W. & Lovell, D. *Proc. 7th ICLAS Symposium, Utrecht*, (Fischer, Stuttgart, in the press).
5. Groen, A. *Lab. Anim.* **11**, 209-214 (1977).
6. Festing, M.F.W. & Totman, P. *Lab. Anim.* 14, 173-177 (1980).
7. Festing, M.F.W. *Nature* **238**, 351-352 (1972).

French museums

SIR — Seen from this side of the Channel, Halstead's criticisms of the new exhibition policy of the British Museum (Natural History) appear extremely unjust (*Nature* **288**, 208; 1980), especially towards the staff who contributed to such clear and remarkable presentation. In France, there is no discussion of that sort, because most of our natural history museums simply have no exhibition policy at all, and are nothing but museums of museums!

Last November, when Halstead was writing his venomous letter, I visited the two new exhibits (on dinosaurs and on the evolution of man) at the British Museum, and was amazed to see how much the public liked them, in particular the children, who seemed very receptive to the logic of cladistic analysis. Halstead's long attack against that method of tracing the interrelationships of living beings deserves more comment. He claims that cladistic analysis, which is the basis of the British Museum exhibits, does not take into consideration the concept of gradualism, according to which evolution can be presented as uninterrupted series of species, older ones being ancestral to younger ones. Hence his accusation against cladism, which is, according to him, trying to make evolution fit Marxist views on the history of societies! Halstead presents gradualism as evidence, and cladism as a crime against evidence. But where is the evidence of gradualism? There is almost none, or rather it is everywhere one wants to see it.

Cladists do not deny that a species can give birth to daughter species, but they claim that in practice, it is impossible to determine the fossil ancestor of a species, since the ancestor is devoid of the derived characteristics of the daughter-species in question. Ancestor–descendant relationship is not, for cladists, a necessary statement, and the history of a group is better expressed by sister-group relationships, often illustrated by a cladogram. Many biostratigraphers are now convinced that successions of species based on stratigraphic distribution are nothing but illusions and that, in reality, each older species is extinguished by the younger one, whose centre of origin may be far from the locality studied. So, the subversive palaeontologists who are supposed to pervade the British Museum, knife between teeth, are simply those who consider that gradualism gives an illusory precision, which is unnecessary for tracing the course of evolution of a group. If these new views are congruent with vague ideological inferences on human societies, so what?. . . After all, the history of human societies often shows a succession of leaps, be they revolutions, wars, epidemics, or changes in climate. But major changes in human societies are seldom decided by the societies themselves. The societies and their chiefs want gradual change (compare the terms "changes in continuity" or "permanent revolution" used by right-wing and left-wing politicians respectively), but in the event, circumstances provoke sudden changes.

In France, some authors are now attacking Darwin (see P. P. Grassé, *L'homme en Accusation*, Albin Michel, Paris) because he is supposed to have inspired Hitler and still inspires the deeply anti-Marxist sociobiologists. Since cladism is also directly descended from Darwin's ideas on phylogeny and systematics, it can just as well be accused of supporting extreme right-wing ideologies.

In sum, even if some cladists claim that this method of analysis is more consistent with their political convictions, it is simply ridiculous to condemn it on the basis of such spurious arguments, or because Halstead's political opinions are different. The British educational system can be proud of the British Museum's exhibitions which will certainly teach future generations of biologists and palaeontologists the principles of phylogenetic reasoning. Perhaps this is just what Halstead is afraid of?

PHILIPPE JANVIER

Laboratoire de Paléontologie des Vertébrés,
Université Paris VI, France

More museums

SIR — Halstead's well publicized reactions[1-6] to the exhibition policies of the British Museum (Natural History), to non-gradualistic hypotheses about the evolutionary process, and to his own mistaken perceptions of cladistic analysis are unfortunate[7-9]. His notion that these views somehow abet non-scientific creationist metaphysics or logically support Marxist dialectics is preposterous. Whether Popper[10], Patterson[11], or Miles (cited by Halstead in ref.4) have said that the evolutionary concept is also metaphysically based and whether this is true are beside the point. Equating Marxist political theory either with cladistic analysis or with so-called macroevolutionary hypotheses is as wrong as social Darwinism was. It needs to be driven home, as Patterson has done[9], that cladistic analysis is about pattern, not about any particular hypothesized evolutionary process, although most phylogenetic systematists do hold that descent with modification is an economical explanation of the existence of biological pattern in general. These systematists also define "related" to mean "genealogically related." In all six publications Halstead confuses cladograms with evolutionary trees. For instance, a steady theme is, "there is no place for ancestral species in a cladogram"[2]. Again, "it is axiomatic, therefore, that no species in the fossil record can be ancestral to any other nor can one species evolve directly into another"[5]. Halstead has not understood that cladograms do not assume that species are ancestors, but neither do they deny possible ancestry when a taxon lacks known derived characters. Assumptions of ancestry are appropriate for trees but are not made in cladograms.

Cladistic analysis parsimoniously estimates relatedness and is therefore testable. The British Museum (Natural History) is to be congratulated for bringing epistemology into its exhibits and teaching visitors that science is a method, not a body of revealed knowledge. That the museum may also need to discuss various hypotheses of evolutionary process[12] and to address other needs as well, such as the public's thirst for discussion of functional anatomy, ecology, behaviour, and so on[13], is also evident.

Malcolm C. McKenna

Department of Vertebrate Palaeontology,
American Museum of Natural History,
New York, USA

1. Halstead, L.B. *Nature* **275**, 683 (1978).
2. Halstead, L.B. *Nature* **276**, 759-760 (1978).
3. Halstead, L.B., White, E.I. & MacIntyre, G.T. *Nature* **277**, 176 (1979).
4. Halstead, L.B. *New Scientist* 17 July 1980, 215-217 (1980).
5. Halstead, L.B. *Nature* **288**, 208 (1980).
6. Halstead, L.B. *Nature* **289**, 106-107 (1981).
7. Patterson, C., Forey, P.L., Greenwood, P.H., Miles, R.S. & Jefferies, R.P.S. *Nature* **277**, 175-176 (1979).
8. Fink, W.L. & Wiley, E.O. *Nature* **280**, 542 (1979).
9. Patterson, C. *Nature* **288**, 430 (1980).
10. Popper, K. *Unended Quest* (Open Court Publishing, La Salle, Illinois, 1976).
11. Patterson, C. *Evolution* (Cornell University Press, 1978).
12. Donovan, D.T. *Nature* **289**, 105-106 (1981).
13. Benton, M.J. *Nature* **289**, 106 (1981).

Sir — I have read with diligence the continuing controversy sparked by Halstead. Thus far I have had difficulty achieving a precise understanding of it. Now, however, certain matters are clear (*Nature* 1/8 January, p.106.) To sum up: Halstead dislikes the new exhibits at the British Museum (Natural History), and he would convince other persons that they, too, should dislike the exhibits. The argument he gives, as far as my evaluation goes, passes not from the ridiculous to the sublime, but emanates entirely from the low end of that spectrum. Knowing Halstead to be usually of good cheer and judgement, I am led to suspect that not all is as it might seem — that the root of his dislike is not to be found brewing in a pot pourri of punctuated equilibria, Marxism, scholastic death, etc. Rather, his dislike may stem, as seems to me, from a sense of loss of "the fossil record" — the ultimate source of the truth of evolution as rendered by a professional class of fellow-specialists. To the dismay, sometimes acute, of the more clerically minded members of this profession, cladistics treats fossils in a secular fashion — not as revelation but as some among many other biological specimens subject to interpretation that is apt, indeed expected, to be diverse, especially with respect to details (for example, the true nature of the "Petralona skull"). As reasonable as this treatment might seem to the outsider, the emotional effect within such a palaeontologist involuntarily confronted with cladistics (as I have witnessed on more occasions than I care to remember) is not unlike that apt to be experienced by a fundamentalist minister who has forced upon him uninvited the notion that the Bible is just one book among many. Suffice it to say that more than one kind of church has been built upon rock.

So what now? Here in the States creationists dislike the museums' secular exhibits on evolution and the schools' secular treatment of that subject. In Britain a palaeontologist dislikes secular exhibits on the "fossil record". The forms are similar, but the substances at first glance seem utterly different. Palaeontology, after all, is nominally a science, and a rational mind can easily defend it as such. The problem I have with Halstead's defence, if I may term it that, is reconciling it with a standard of rationality.

Gareth Nelson

Department of Ichthyology,
American Museum of Natural History,
New York, USA

Genes and racism

Sir — Steven Rose notes in his recent letter (*Nature* 22 January, p.335) that a National Front journal *New Nation* has claimed to find support for racism in my writings on sociobiology, as well as in those of Dawkins and Maynard Smith. Rose calls on the latter two authors to dissociate themselves from such misuse, although curiously he does not extend the same invitation to me. To keep the record straight, I am happy to point out that no justification for racism is to be found in the truly scientific study of the biological basis of social behaviour. As I stated in *On Human Nature,* "I will go further and suggest that hope and pride and not despair are the ultimate legacy of genetic diversity, because we are a single species, not two or more, one great breeding system through which genes flow and mix in each generation. Because of that flux, mankind viewed over many generations shares a single human nature within which relatively minor hereditary influences recycle through ever changing patterns, between the sexes and across families and entire populations".

If there is a possible hereditary tendency to acquire xenophobia and nationalist feelings, it is a *non sequitur* to interpret such a hypothesis as an argument in favour of racist ideology. It is more reasonable to assume that a knowledge of such a hereditary basis can lead to the circumvention of destructive behaviour such as racism, just as a knowledge of the hereditary basis of haemoglobin chemistry and insulin production can lead to the amelioration of their pathological variants.

I now call on Professor Rose to consider these and similar arguments raised in my writings. It is my hope that he will not confine himself, as he has in the past, to arguments that link sociobiology to racism and thus to continue to abet the very misuse which he piously claims to deplore.

Edward O. Wilson

Harvard University, Cambridge,
Massachusetts, USA

Origin of cancer

Sir —John Cairns' article on "The origin of human cancer" (*Nature*, 29 January 1981, pages 353-357) dismisses the importance of chemical mutagens in human cancer aetiology on, we believe, very tenuous grounds. He argues that there must be a single underlying mechanism for tumour production and that this is not through point mutation. As your leading article (*Nature*, 5 February 1981) quite rightly points out, why should there be a single mechanism for cancer development? Indeed, can one say that cancer is even a single disease? Cairns' article appears to us to be a simplistic approach to a complex problem. We all would like to have the answer to how a cancer cell develops, but can one say that Cairns' article will lead us any nearer to the truth? Why should workers in the field of chemical carcinogenesis abandon a large body of work which may be getting somewhere near to establishing why populations differ in cancer incidence, and substitute a hypothesis of genetic transposition bearing in mind that, as Cairns writes, it is not yet clear whether transposition is important in vertebrate development. Should we abandon the somatic mutation theory of cancer for a hypothesis with no experimental evidence? Furthermore, why should mutagens/carcinogens only act through point mutations? It is well recognized, as Cairns states, that agents reacting with mammalian cell DNA can cause gross effects such as chromosome aberrations, sister chromatid exchanges, deletions, and so on. Just because one measures DNA reaction in bacteria with a point mutation system does not mean *a priori* that this is the mechanism of action of these chemicals in mammalian cells. After all, bacteria don't even get cancer. Bacteria are merely used for carcinogen screening because they are cheap and mutations are easy to score.

Where Cairns' article is mischievous is in suggesting that people are wasting their time looking for carcinogens in the environment and then, having identified them, seeking to reduce exposure. The fact is that if people would stop exposing themselves and their immediate families to mutagens/carcinogens in the form of cigarette smoke, a large proportion of cancers would be prevented (besides a large proportion of coronary heart disease, lung disorders, etc.). If Cairns accepts that smoking is bad for one, why should he not accept that other environmental insults might also be carcinogenic? To say that animal carcinogens induce mainly liver cancer, that humans don't normally get cancer of this organ, and therefore animal liver carcinogens have nothing to do with human cancer shows a complete unawareness of chemical carcinogenesis. Most chemicals which have been identified as human carcinogens do *not* give the same spectrum of tumours in animals as in humans, for example benzidine causes bladder cancer in man but liver cancer in animals.

Cairns' main argument for DNA reaction being unimportant in carcinogenesis is the finding that in xeroderma pigmentosum patients, few if any internal cancers have been seen. Why should one expect an increased incidence of lung cancer, for example, in these patients? Is the skin the same as the lung in its function, biochemistry, enzyme profile, etc.? Why should we expect the mechanism of skin cancer to be the same as for internal organs? Do xeroderma pigmentosum patients live in the same environment as the normal population? I don't think one has to indulge in "special pleading" to support the case of somatic mutation as being important in cancer production. The majority of facts available tend to support the somatic mutation hypothesis. When there are sufficient data available to overthrow this hypothesis, then is the time for it to be abandoned. In the meantime, those of us working in chemical carcinogenesis will carry on identifying carcinogens in the environment and recommending that exposure be reduced. Whether society (or Cairns) listens to us is entirely up to them.

R. Colin Garner
Paul J. Hertzog

Cancer Research Unit,
University of York, York, UK

The Great Badger Debate

What follows is a brief summary of the report Badgers, Cattle and Tuberculosis, *published by the British Ministry of Agriculture, Fisheries and Food last October and a comment by its author, Lord Zuckerman, on various criticisms that have been raised, some of them in* Nature

Report and implications

Bovine tuberculosis has been recognized as a serious problem in the United Kingdom (and also a source of human tuberculous infection) at least since 1913, when steps were first taken to control the disease. Farmers were encouraged to register herds free from tuberculosis by means of a premium price for the milk they produced, and infected cattle were also slaughtered compulsorily. But bovine infection continued, with unexplained "breakdowns" even in attested herds.

The link between bovine tuberculosis and badgers was first recognized in 1971, with the discovery of tuberculous lesions in a badger found dead. The organism responsible was found to be *Mycobacterium bovis*, that responsible for bovine tuberculosis. It now appears to be common ground that badgers can be infected by this organism. In 1975, ministry officials began controlling badgers found within a specified radius of infected herds of cattle by the use of cyanide gas. These procedures were abandoned in October 1979, but have been resumed in South West England since the publication of the Zuckerman report.

The transmission of tuberculosis from badgers to cattle is accepted in the report, which says that viable *M. bovis* have been recovered from grass sprayed with the urine of infected badgers (but more easily in the winter than the summer). The report also says that during the period in which control by gassing was carried out, "185 out of 2,432 badgers that had been collected in the vicinity of 477 farms where breakdowns had occurred" turned out to be infected with *M. bovis*. The report also records how, in New Zealand, the opossum was discovered to be a reservoir of bovine tuberculosis, and gives (in its Appendix III) accounts of how, at several farms in South West England, infection in badgers has been linked with breakdown among herds of cattle.

The South West of England sustains a third of the United Kingdom's cattle, and also dense populations of badgers. It is common ground between Lord Zuckerman and those who have criticized his report that badgers are also prevalent in parts of England, where bovine tuberculosis is rare. Lord Zuckerman concludes, in his report, that the population density of badgers determines both their susceptibility to infection and the risk of transmission to cattle. The frequency of infection among badgers in the South West of England has been found to be as great as one in five in some places.

Among the recommendations in the report were the resumption of badger control by gassing in the South West, the investigation of the incidence of tuberculosis among badgers in contiguous parts of the country, financial support for field studies of the British badger population, the publication of annual reports of the progress of the control campaign and a more thorough review of what will by then have been learned of the badger problem after an interval of three years.

Zuckerman's case

The 6 November 1980 issue of *Nature* (p.6) carried a brief account of a report on tuberculosis (TB) in badgers and cattle which I had prepared for the Minister of Agriculture, Fisheries and Food. The report was written in response to a request for an assessment of the soundness or otherwise of the scientific evidence on which the ministry based the view that badgers in parts of the South West of England constitute a reservoir of the bovine tubercle bacillus. Because of its statutory responsibility to suppress TB in cattle, I was also asked whether the government was justified in its policy of destroying badgers where there was presumptive evidence of the presence of tuberculous badgers that might be responsible for outbreaks of TB in neighbouring cattle.

During my investigation I had consulted several scientists engaged in the study of animal disease, as well as practising veterinary surgeons. All agreed with the findings of the ministry's scientific and veterinary officers with many of whom I had conferred. My own review of the evidence left me satisfied that badgers in certain areas of the South West are heavily infected with the bovine tubercle bacillus, and sometimes transmit the disease to cattle grazing on pasture that has been contaminated with their sputum, pus, urine and faeces. I uncovered not a vestige of scientific foundation for widely publicized statements from "protestors" to the effect that none of the story was "true" and that, in effect, the whole thing was a hoax perpetrated by government servants.

It was, of course, unlikely that a fresh display of the facts would necessarily succeed in "converting" the opposition. But it came as a surprise to read first, in an unsigned article in *New Scientist* (4 December 1980, p.619), and then in a letter published in *Nature* intended as a corporate statement by The Mammal Society (11 December 1980, p.532), that the conclusions of my report, in which I had assembled not only more, but also more up-to-date, information on the subject than had been made available before, were questionable on scientific grounds. Apart from agreeing that "the badger is a major reservoir of bovine TB in certain limited areas of South West England, and hence a potential danger to the cattle in those areas", the society asserts that my report gave a "biased interpretation of the evidence", that "many of (my) conclusions were not justified" by the data which I had presented, and that "factually misleading statements" therefore needed to be corrected.

The report

My main conclusions, set out in the report, were first, that "by any rational epidemiological standards, badgers (in some parts of South West England) now constitute a significant reservoir of the bovine strain of the tubercle bacillus". The Mammal Society thinks otherwise. It declares that "there is no scientific evidence to justify" the view that badgers in the affected areas constitute "a highly infected population". If the society is sufficiently well informed to make this extraordinary claim, the onus of providing proof, as Dr Plowright has already pointed out (*Nature* 1/8 January 1981, p.8), is on them.

Next, the Mammal Society was, to say the least, irresponsible in their quotation of what was in the report. I did not "overlook" the fact that during the first seven months of the moratorium the percentage of badgers with TB had fallen slightly in Cornwall at the same time as it had increased in Gloucestershire and Avon; this is precisely what my figures revealed. Nor did I interpret these or any of the other available figures in the "black or white" way that the Mammal Society suggests I did and, now in reverse, that it itself does. The interpretations I provided were the best that could be put on the figures, and were carefully qualified. Conclusions contrary to those I drew from the facts would have been absurd.

The Mammal Society is not criticizing when it merely repeats what I emphasized: namely, that the disease in badgers occurs only in "pockets". But it is falsifying what I did write when it says that I had asserted that badgers in the affected areas of the South West are spreading TB to badgers in other areas. What I wrote was that no one knows how one pocket of infection in badgers relates to another (paragraph 134 of the report). And I did not say that TB *is* (my italics) a major hazard to the survival

0028-0836/81/080628-03$01.00

of the badger (even if it clearly is to the badgers in the areas already affected). What I said was, were the disease "to take hold in badgers in other parts of the country" "there is no saying what the consequences would be . . . as they relate to the survival of the badger".

In considering the persistence or extinction of a disease in a given locality, the statistical laws governing its spread must be borne in mind. If the average number, x, of new infections from a unit which becomes infected is less than 1, the disease will die out. If greater than 1, the disease, once established, will spread. However, if x is not markedly greater than 1, there is still a good chance that a single infection will not result in its establishment; but if repeated new infections occur, then eventually one of them will be successful in establishing the disease.

With badgers two separate units are involved: first, the individual within a badger sett, and second, the setts themselves. As Dr Plowright has re-emphasized, the epidemiological conditions are highly favourable for transmission of TB within a sett. Transmission between setts seems to depend very greatly on such factors as density of population, contiguity of setts, and fighting between members of different setts. In an area which adjoins another where the disease is endemic, there will be a continuous flow of infection from the infected area to the border setts of the other area. But if x for the second area is less than 1, the infection will not spread through the whole area.

Infectivity

It is obvious that the period during which a unit remains infective depends on the nature of the disease. For a disease which kills quickly or in which there is complete recovery without the unit becoming a carrier, the period will be short. In a disease such as TB it will be considerably longer. Note also that the value of x will change with time, both because of the direct effect of the disease on population density and of acquired immunity, and also because of other unrelated factors.

I have discussed these matters with Dr Frank Yates of Rothamsted. In his view, my supposition that bovine TB may have been prevalent in badgers throughout the country before its virtual eradication from cattle is almost certainly justified (paragraph 133 of my report). After its eradication from cattle and the consequent absence of cross-infection the disease died out in areas of the country in which x was less than 1, as it seems to have done in most of the UK.

Dr Yates has also pointed out that once the disease has died out in an area, this area will remain free as long as no new infection is introduced, but if owing to an increase in the density of population or for other reasons x becomes greater than 1, the disease, once established, will then spread until the whole area is infected. The question put by the Mammal Society as to why areas in other parts of the United Kingdom with a badger population density which is presumed to be as great as in the South West were not infected, seems basically simple. Either x has remained less than 1, or they have just been lucky.

Before turning to certain further information about the prevalence of TB in badgers that has become available since my report was written, it is necessary to refer to what I said about the incidence of breakdowns in cattle, since this is also something about which the Mammal Society is critical.

The campaign to suppress bovine TB resulted in a decline in the incidence of reactors in cattle herds, not only in the South West, but also in the rest of England. This the Mammal Society acknowledges. But, as Dr Plowright has already pointed out the society failed to note in its statement the critical fact that while the incidence of herd breakdowns has fallen dramatically in all parts of the country once the culling of tuberculous cattle was made compulsory in 1950, in the affected areas of the South West it has consistently remained more than five times the national average. This difference was the essential reason for the search for a "cause", and for the resultant finding of a reservoir of the bovine tubercle bacillus in badgers in the affected areas. And here is where the Mammal Society's criticisms have gone wildly astray. Their suggestion that one possible interpretation of the figures given in the Report was that unknown "subtle factors" resulted in a decline of TB in both cattle and badgers throughout the country, and that gassing had little significant effect on the overall timing or rate of decline is, as Dr Yates has shown (*Nature* 22 January 1981, p.218), based on a very superficial and distorted use of the evidence. His more detached presentation clearly indicates that the gassing campaign appears to have had a considerable effect in reducing the very high incidence of herd breakdowns in the worst affected parts of the South West. What the Mammal Society did was lump together in a graph my figures for TB in badgers in all the affected counties of the South West, whereas as Dr Yates has shown, there is a radical difference when one compares the parallelism in the prevalence of TB in badgers and the incidence of breakdown in cattle as between Cornwall and Gloucestershire on the one hand, and Devon, Dorset, Somerset and the rest of England on the other.

In its haste to criticize, the Mammal Society has also clearly failed to realize that the data assembled in my report were the best available at the time, and that more would be published as they became available — I have recommended annual reviews. The society's charge that it was "manifestly untrue" that herd breakdowns reflected a high local prevalance of TB in badgers was based on the figures for Cornwall, where only 15 per cent (51/340) of the breakdowns were attributed to badgers. What the society failed to bring to the attention of its readers — their examination of the evidence must indeed have been cursory if they themselves failed to notice it — is that, as Dr Yates did note, in Cornwall 69 per cent (234/340) of the breakdowns were classed as unknown, whereas in the rest of the South West (plus Sussex) only 26 per cent (107/417) were so classed, 67 per cent (280/417) being attributed to badgers.

Lord Zuckerman (right) with Agriculture Minister Peter Walker and the report

Figures

Before rushing into print the society might indeed have enquired further about the Cornish figures. I have now obtained a breakdown by years which gives:

Cornwall	*Unknown*	*Badgers*
1972–73	110	–
1974–75	88	10
1976–78	36	41
	234	51

The explanation is simple. Badger investigations were only begun in Cornwall as a routine measure in 1974. The progressive reduction in the proportion of "unknowns" indicates their increasing thoroughness. I am also informed that for the whole of Cornwall 72 per cent of the 51 outbreaks attributed to badgers were within 2 miles of setts containing infected badgers, and 33 per cent were within half a mile. For the West Penwith area the corresponding percentages were 97 and 47.

The data which I had when I wrote my report related to only the first seven months of the period of the moratorium (Table 18 of my report). I have now been provided with figures for the first twelve months during which no new gassing operations were carried out. During the year, 611 badgers were autopsied in connection with official investigations to determine the source of herd breakdowns in the affected areas of the South West, and 73 (12 per cent)were found to be infected. Needless to say, there were differences between the counties concerned, but since the significance of these differences could only be determined through a detailed analysis aimed at revealing their relation to specific herd breakdowns, they hardly merit discussion.

More important is the information that is now available about the herd breakdowns that occurred in the affected areas in the first twelve months of the ban, about which nothing positive could be said on the basis of the data that were made available to me when I drew up my report, and in which reactors were found in only 0.7 per cent of the 17,411 herds that were tested. Between April and September a further 6,845 herds have been examined, bringing the total up to 24,256. The percentage of reactor herds had risen substantially in comparison with the first seven months of the moratorium in every one of the affected counties of the South West. Overall, the percentage for the period April to September was 2.2, double that for the period January to September 1979. For the year of the moratorium as a whole it was some ten times higher than for the rest of England.

The anonymous article in *New Scientist* (4 December 1980) referred not only to the Mammal Society's statement, but also (1) to certain critical comments by Dr Hans Kruuk that follow the same line as the society's statement; and (2) to an assertion by Miss Eunice Overend, described as an "enthusiast" of badgers, about the speed with which TB develops in badgers; she is quoted as saying that while TB may develop quickly in captive badgers, "it wouldn't happen in the wild". In answer to an enquiry I put to her, she also challenges the views of veterinarians about the speed with which the disease can take hold in cattle.

I have no doubts about Miss Overend's sincerity, but her "facts" happen to be wrong. Advanced TB lesions have been found in the lungs of wild badger cubs ranging in age from four to nine months, which had clearly been infected by routes other than bite wounds, and it is estimated that at least some of these cubs would have succumbed to the disease in a matter of weeks.

Anomalies

Of course, there are still many "anomalies" in the story, which further enquiry might elucidate. But every one of the so-called anomalies referred to in the Mammal Society's manifesto, all its speculative questions, "why this, why not that", are, as Dr Plowright has pointed out, mentioned in my report. Setting them out as a form of criticism adds nothing to the story. I hope that the further enquiries which I suggested should be undertaken to clear up a number of "unknowns", and which the Minister of Agriculture, Fisheries and Food, in his statement to Parliament said would be started as soon as the necessary arrangements can be made, will be entrusted to experienced and competent hands.

So much for the "scientific" criticisms that have been levelled at my report. No new facts have been brought forward. Nothing that has been said would make it anything but irresponsible for the government to devise a policy of action for the suppression of bovine TB, to which it is statutorily committed, other than what it is doing. Unfortunately, the Mammal Society's statement has been used as part authority for a press-release by the UK branch of the World Wildlife Fund, in which the government's acceptance of my report is in effect condemed. This was issued without any consultation with the Animal Health division of the Ministry of Agriculture, Fisheries and Food (MAFF).

As a scientist I have been a dedicated supporter of the conservation movement since before the Second World War,and as a scientific adviser to government I played a part in organizing our own formal institutions for the protection of the environment. I therefore find it regrettable that the only way I could have found favour with those who are concerned to obstruct the government's anti-bovine TB policy would have been either to suppress, or to fudge, the evidence which had been gathered by those scientists primarily concerned over many years with the problem of TB, not only in cattle, but also in badgers (as well as in those luckier representatives of the wild fauna of the country which, by comparison, are relatively resistant to the infection).

Mammal Society

The Mammal Society was one of the bodies which submitted written evidence to help my enquiry, citing Stephen Harris and Dr Kruuk as its two authorities since, so the Society said, they were "the two principal non-MAFF scientific researchers on badgers in this country". Presumably both helped in the drafting of the corporate statement which was designed to correct my "factually misleading statements". Since there can be few land-mammals that have generated as much proprietary interest as has the badger among the handful of naturalists who have added to our knowledge of the ways of the animal, I should note that I cannot find in the zoological literature any paper by Harris on badgers. Dr Kruuk's present studies relate to the animal's natural history, and I referred (paragraph 141 of my report) to a proposal which he had already published, that "pastures in infected areas should be so dressed as to reduce drastically the population of earthworms which they sustain", commenting, however, that I did not know "whether there is a way of doing this without detriment to the cattle which also forage the pastures that would be treated", nor "what those concerned with the welfare of birds — which also eat worms — would think about the measure which Dr Kruuk has proposed" Before turning to the badger, Dr Kruuk studied the behaviour and ecology of gulls and carnivores — mainly hyenas. For this work he was awarded the Zoological Society's Scientific Medal in 1975.

The errors in the Mammal Society's statement suggest that its badger panel did not include experienced epidemiologists or veterinary scientists. That it should have done is clear. For contrary to what has been said to me by the society's president, Dr Ernest Neal, the study of mammals is not just "a rather narrow field of interest" that can be covered by watching the behaviour of animals in the wild. Whether or not it was intended, the misleading impression which the statement has made on the media, and on those who now clamour against the government's decision, was that the Mammal Society is a body that can make pronouncements on animal disease which, if anything, might have been expected from the Royal College of Veterinary Surgeons, or from a commitee set up *ad hoc* by the Royal Society.

Of course members of the Mammal Society are free if they so wish to criticize the scientific validity of the data that I collected and to question my interpretations, but by insisting that it has the right to make a corporate, but anonymous, statement of the kind it has the young Mammal Society has, in my view, done itself harm. Here it is useful to refer to the main editorial in *Nature* of 11 December 1980 (p.525) entitled "Code of conduct for national academies", which pointed out that even "The Royal Society has no means by which it can reach a corporate opinion on any substantial issue of public policy . . . for example, say, government policy on nuclear power stations or whether manufacturers of drugs should in future be liable for side effects". Of course, the Royal Society could set up committees to look into such matters, but then its members would be named, and those Fellows of the Society not on the committees would certainly not be committed to whatever recommendations were put forward. If it wishes its published statement to be considered further then the Mammal Society should clearly name the "experts" who were responsible for its drafting. Dr Yates suggests that what the Mammal Society has done illustrates "the way in which distorted and partial graphic presentation can be used for propaganda purposes". It might have been more charitable to say that what was done was done out of ignorance. But then, as Gunnar Myrdal has put it, one has to remember that ignorance, like knowledge, can be steered for a purpose.

Not only was there no scientific value in the Mammal Society's statement; it was clearly the work of only a handful of its members, none of whom, so far as anyone can tell, had ever worked on the epidemiology of tuberculosis in animals. It is unfortunate that the society's action should have encouraged those who protest against MAFF's findings and policies in the belief that they have scientific support for the new wave of opposition which they have already mounted.

NEWS AND VIEWS

Too soon for the rehabilitation of Lamarck

CAN acquired characteristics be inherited? The arresting letter from Gorczynski and Steele (on page 678 of this issue) says that there are circumstances in which they can be. The interest of the observations is beyond dispute. The days have long since gone when dark lamarckian issues such as these could be raised respectably. Even in primary schools, or at least in those not yet captured by the creationists, young people are now told that the giraffe did not come by its long neck because earlier generations of giraffes had stretched their necks in an attempt to eat the leaves of ever-taller trees, and had been able to pass on the characteristics thus acquired. So how can it be that dark lamarckian evidence can be uncovered at this late stage in the study of inheritance? That is what many people will be asking. The simple answer is that it is too soon to know what to make of the observations now reported. On such an important issue, no single set of observations can be sufficient to overturn the accepted view of how inheritance is organized in living things. Independent confirmation is now the most urgent need. So much has been clear since the first article on this subject by Gorczynski and Steele a year ago (see R.B. Taylor *Nature* **286**, 837; 1980). By all accounts, confirmatory evidence has been hard to come by. Some reports of failures to confirm seem destined to appear in the next few months. Gorczynski and Steele, like their would-be critics, have chosen an intricate field in which to work. For the time being, however, the results they report must be put firmly in limbo. They do not constitute a proof that Lamarck has been right all this past century and a half, but only a question to be pondered.

The lamarckian heterodoxy, it must be remembered, had respectable beginnings. Lamarck had the wit to recognize the need for an explanation both of the diversity of species and the evidence for evolutionary processes uncovered (by the Lyells of this world) from the sedimentary rocks. At the outset, the notion that characteristics once acquired might be inherited cannot have failed to be seductive. What could have seemed more sensible, in the first half of the nineteenth century, than that a giraffe with the good sense to stretch its neck in the search for better food might produce offspring with similar characteristics? The difficulty, from the beginning of the Lamarckian thesis, is that it has never amounted to an explanation. If the giraffe came by its long neck thanks to the diligence of its ancestors, why does not a woman whose arm is amputated produce one-armed children? Lamarckians, both ancient and modern, have not been able to answer such questions.

Historically, none of this has mattered very much. Darwin's explanation of both diversity and the appearance of change, or evolution, turned out to be a better explanation. Both diversity and change, said Darwin, could be explained by the mechanism of evolution by natural selection. Longer-necked giraffes might be fitter in the sense of being the more able to propagate themselves. So much is now accepted, yet darwinian scholars are forever gleefully discovering in *The Origin of Species* or other darwinian texts evidence of Lamarckism in Darwin himself. So what? Formally, Lamarckism and Darwinism are not mutually exclusive. One does not disprove the other. Although Darwin showed that natural selection is the predominant process in evolution, the data available in the second half of the nineteenth century could not have excluded the admixture of a modicum of lamarckian inheritance. The contemporary orthodoxy stems not from Darwin but from the turn of the century. The rediscovery of mendelism provided a mechanism by which variations could arise and thus be selected. Weissman's doctorine that there is a sharp distinction between the cells of the germ line and the soma meanwhile outlawed Lamarck. In modern language, according to Weissman, the genotype has a decisive influence on the phenotype, but the phenotype can affect the genotype only because the phenotypically fitter individuals are more adept at the propagation of their own kind. So far, everything that has been learned about the mechanisms of ineritance in animals and plants is consistent with both Darwin and Weissman.

Laboratory experiments in which viruses have been shown to carry new information into the germ line are nevertheless a tantalizing hint that there may be exceptions to the rule. Gorczynski and Steele challenge Weissman's doctrine, not the belief that darwinian selection is predominant in the evolution of species. Their latest experiments are an extension of those described a year ago. Then, they claimed that male mice made tolerant to a single histocompatibility antigen produced, from normal mice of the same strain, offspring that were also tolerant of the same antigen, at least in the first and second generations. The acquired characteristic (tolerance of a specific antigen) appears to be inherited. The latest series of experiments begins with male mice made tolerant of two antigens. Again, the data seem to suggest, tolerance to both antigens appears in the first and second generations. On the evidence available — and the numbers are only small — the two acquired characteristics are transmitted independently to the first and second generations of mice. Gorczynski and Steele put forward a possible explanation of their results. Weissman's doctrine has been violated, they say, because some naturally occurring RNA virus may have picked up the altered genes from the immunological system of the mice made tolerant to histocompatibility genes, and may have transferred these to the germ line. The results, they say, "demand an explanation not offered by conventional neo-darwinian genetics".

Not everybody will agree. For one thing, the experiments so far described are incomplete. If, indeed, a somatic mutation has been transferred from the soma to the germ line, then it should be possible to select both male and female mice of the first generation carrying the gene for tolerance and demonstrate that, thereafter, the mice breed true. In reality, the frequency of tolerant individuals is less in the second generation than the first, suggesting that whatever is being inherited cannot be fixed in the germ line. Moreover, there are other possible explanations for the results so far reported. One suggested by R.B. Taylor is that the mice may have inherited not the gene for tolerance but a gene for the production of the antigen used for the induction of tolerance. Other explanations will emerge. While so little is known of the mechanism of tolerance in these mice, and while conventional explanations have not been ruled out, it is premature to call for a recasting of genetics, "neo-darwinian" or otherwise. And the experiments need to be repeated.

Gorczynski and Steele will no doubt

0028-0836/81/080631-02$01.00

regard this response to their interesting article as toffee-nosed. Others, sadly, will go further and say that even a single case of the inheritance of an acquired characteristic implies that contemporary genetics is a house of cards. That is an outrageous leap ahead.

Even if the data now reported are confirmed, the explanation of them offered by Gorczynski and Steele may turn out to be no more than an exception that proves the conventional rule. For if it requires the extreme procedures that have been used in these experiments to provoke such an indecisive example of lamarckian inheritance, may not the generality of Weissman's doctrine be strengthened? Even that speculation is however premature. It will be time enough to worry about the rehabilitation of Lamarck when and if the data now reported are independently confirmed. □

Gamma-ray astronomy

from R. D. Wills

THE spectral range of gamma-ray astronomy is very broad. At a recent meeting* results were presented and discussed which span seven decades of photon energy from about 10^5 to more than 10^{11} eV. The techniques described differ considerably along the energy range, but essentially two approaches are adopted to overcome the problem of atmospheric absorption. At the lower energies, where the fluxes are relatively higher, instrumentation is carried above the atmosphere by balloons or satellites. Above about 10 GeV it is advantageous to measure at sea level or mountain altitude the secondary particles generated by interactions of the primary radiation in the upper atmosphere. These secondaries can be detected over a wide area enabling very low fluxes to be measured.

Most of our knowledge about high-energy (30 MeV to 5 GeV) gamma rays has been obtained from the NASA satellite SAS-2 (C. E. Fichtel *et al. Astrophys. J.* **198,** 163; 1975) and its ESA successor COS-B (L. Scarsi *et al. Proc. 12th ESLAB Symp.* ESA *SP-124,* 3; 1977). The latter has made the first complete detailed survey of the Milky Way in high-energy gamma rays (H. A. Mayer-Hasselwander *et al. Ann. N. Y. Acad. Sci.* **336,** 211; 1980). On behalf of the Caravane Collaboration, H. A. Mayer-Hasselwander (Max-Planck-Institut für Extraterrestrische Physik, Garching bei München) reported that, up to the end of 1978, 83,233 gamma rays had been recorded from the galactic disk (latitudes $|b| < 12°$) and outlined the current analysis of their spatial distribution and energy spectrum. A. W. Strong and A. W. Wolfendale (University of Durham) reviewed the results at intermediate latitudes ($10° < |b| < 20°$). As in the lower-latitude range, the gamma-ray intensity is well correlated with the various tracers of the constituents of interstellar matter. The use of galaxy counts as a tracer of the 'total gas' enabled the gamma-ray emissivity of the interstellar medium to be obtained. Its energy dependence showed clearly the importance of the contribution of cosmic-ray electrons to the generation of gamma rays. G. F. Bignami (Istituto di Fisica Cosmica del CNR, Milan) drew attention to low-latitude correlations with galactic structures on a variety of scales, and pointed out that inconsistencies between the gamma-ray map and Gould's Belt could be explained in terms of the newly discovered 'Dolidze Belt' (M. V. Dolidze *Sov. astr. Lett.* **6,** 51; 1980).

Several speakers discussed the localized gamma-ray sources, of which 25 are listed in the 'Second COS-B Catalog' (B. N. Swanenburg *et al. Astrophys. J.* in the press). P. A. Caraveo (Istituto di Fisica Cosmica del CNR) described a search for X-ray counterparts in the error boxes of some of the unidentified sources, whose positions are known only within a radius of about 1 degree. On the basis of their narrow latitude distribution R. Buccheri (University of Palermo) proposed that at least a significant part of them could (like the only two unambiguously identified sources, PSR 0531 + 21 and PSR 0833–45) be very young pulsars, but ones not yet known to radioastronomers. A complementary suggestion by T. Montmerle (Centre d'Etudes Nucléaires de Saclay) was that a class of gamma-ray sources is extended sources, associated with selected active regions in the Galaxy, like the ϱ Oph cloud or the Carina complex.

Extragalactic gamma rays were reviewed by D. Ramsden and A. J. Dean (University of Southampton). In the high-energy range only the quasar 3C273 has been detected (G. F. Bignmai *et al. Astr. Astrophys.* in the press), though upper limits exist for the fluxes from many active galaxies and some galaxies in the local group (G. F. Bignami *et al. Astrophys. J.* **232,** 649; 1979; A. M. T. Pollock *et al. Astr. Astrophys.* in the press). The Seyfert galaxy NGC 4151 has been detected below 20 MeV in balloon experiments (F. Perotti *et al. Nature* **282,** 484; 1979) but lack of confirmation from other experiments indicates that it may be variable at these energies, as it is in other parts of the electromagnetic spectrum. In this energy range it is possible to detect gamma-ray line emission due to nuclear interactions or antiparticle annihilation (R. Ramaty and R. E. Lingenfelter *Nature* **278,** 127; 1979). The experimental status of such measurements was reviewed by L. E. Peterson (University of California, San Diego), while a theoretical overview was given by R. Ramaty (NASA/Goddard Space Flight Center). Both speakers mentioned the detection of spectral lines in transient (burst) events. Those of 19 November 1978 and 5 May 1979 exhibited a feature at 0.4 MeV which is believed to be gravitationally redshifted positron-annihilation radiation, probably from the vicinity of a neutron star. G. Vedrenne (Centre d'Etude Spatiale des Rayonnements, Toulouse) described how the accuracy of locating bursts by triangulation had been improved (down to 1 arc min), using the long baselines available from interplanetary spacecraft such as Helios, Pioneer Venus, Venera and ISEE-3. However, no firm conclusion on the nature of burst sources could be drawn from the distribution of their locations, nor from the burst size (log N — log S) distribution.

Gamma rays of energy greater than 10^{11} eV can be detected at ground level by virtue of the Cherenkov light from the small air showers they generate in the Earth's atmosphere. K. E. Turver (University of Durham) and T. C. Weekes (Smithsonian Institution, Mount Hopkins Observatory) described recent developments in this field and emphasized the possibility of extending the technique to lower energies using multi-element detector arrays. Of the sources already detected, new results on Cygnus X-3 were presented by A. A. Stepanian (Crimean Astrophysical Observatory) and N. A. Porter (University College, Dublin). Reviewing investigations of pulsars, B. V. Sreekantan (Tata Institute for Fundamental Research, Bombay) remarked especially on the results from Ootacamund, where pulsed emission from PSR 0531 + 21 had been detected in February 1977, but despite increased sensitivity, not in the succeeding two winters (B. N. Bhat *et al. Astr. Astrophys.* **81,** L3; 1980). There was some speculation as to whether this variability might be related to a change in the 50–3,000-MeV light curve which COS-B had detected between 1975 and 1979 (presented on behalf of the Caravane Collaboration by R. D. Wills, Space Science Department of ESA).

Future prospects for gamma-ray astronomy were reviewed by C. E. Fichtel (NASA/Goddard Space Flight Center). In the next decade satellites such as the Soviet/French 'Gamma-1' and NASA's

---

*The Royal Society's meeting for discussion on 'Gamma-Ray Astronomy', arranged by the British National Committee on Space Research, was held in London on 27–28 November 1980. The proceedings will be published in the *Philosophical Transactions of the Royal Society.*

R. D. Wills is in the Space Science Department of the European Space Agency Space Research and Technology Centre, Noordwijk, The Netherlands.

0028-0836/81/080632-02$01.00

Gamma-Ray Observatory should carry experiments which will have over an order of magnitude greater sensitivity than those flown so far while in some parts of the energy range substantially improved energy and angular resolution will be achieved. Such experiments should greatly enhance our knowledge of several astrophysical phenomena especially those associated with compact objects, astrophysical nucleosynthesis, galactic structure, the origin of cosmic rays, active galaxies and universal matter–antimatter symmetry. Meanwhile the meeting learned from R. D. Wills (Space Science Department of ESA) that COS-B was still recording good data after more than 5 years in orbit and that ESA's Science Programme Committee had recommended the continuation of its mission into 1981. □

Diffusionless reactions and crystal engineering

from J. M. Thomas

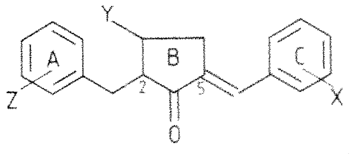
Diffusion, which is the spontaneous kinetic drift of atoms, ions and molecules, is a phenomenon so inextricably mingled with a wide range of chemical phenomena that it is difficult to imagine that there could be a class of chemical transformation in circumstances where even the idea of diffusion is superfluous. The migration of ions in solution or in solids, the dissolution of an electrode in an electrolytic cell, the transport of reactants to, and products from the surface of a heterogeneous catalyst are all well known examples of chemical processes in which diffusion plays a dominant role. Indeed there are many kinetic situations — the recombination of iodine atoms in the gas phase is a renowned example — where the diffusion of the reactants is rate-limiting. Under which set of circumstances, therefore, is diffusion of such little consequence?

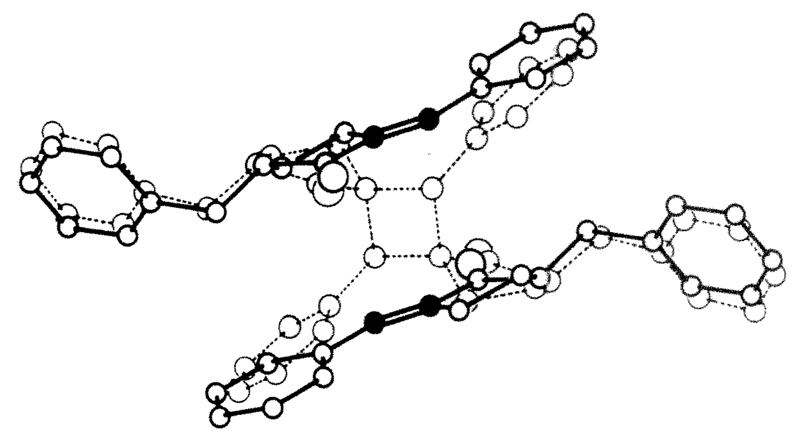
It is now clear that very many highly efficient and selective chemical reactions may proceed smoothly over a time scale of minutes or hours *in the solid state* without the need for prior diffusive motion. So frequently has it been remarked that chemical reactions are sluggish in solids because diffusion in the solid state is itself a very slow process, that it often comes as a surprise to discover that crystals of certain organic molecules can be induced to react rapidly, and in a preordained fashion, even at cryogenic extremes. The secret is to design these organic molecular crystals — to 'engineer' them — in such a way as to pack together potentially reactive monomer units which can subsequently be 'compelled' to undergo preferred reactions, which entail no diffusion but merely minor reorganization — a slight rotation or a subtle shift of part or all of the reactant molecules. Such reactions, which are termed topochemical, yield clean products, the chemical nature and crystallinity of which are governed more by the molecular environment within the parent crystal than by the intrinsic reactivity of the molecule.

Crystal engineering consists essentially of selecting molecules of a given class and functionality according to their size, shape and stereochemistry. They are then 'persuaded' to crystallize into certain desired crystallographic space groups; and within those crystals, by appropriate stimulation — UV irradiation, for example — diffusionless, intermolecular reactions are allowed to proceed.

Many aspects of this new approach to synthetic chemistry are an echo of the past. Kohlschutter (*Z. anorg. allg. Chem.* **105**, 121; 1918), the first to use the term topochemistry, realized that certain host crystals, notably graphite, dictate the degree of topotaxy associated with a solid-state reaction. Topotaxy refers to the mutual relationship between the unit cell axes of the daughter and parent crystals — in this case the graphite intercalates (for example, graphite–ferric chloride or potassium graphite) and the graphite, respectively. The late Gerhardt Schmidt, who was amongst the first to talk of crystal engineering, along with his associates at the Weizmann Institute, made many noteworthy contributions to this field (see Schmidt *Pure appl. Chem.* **27**, 647; 1971; Cohen and Green *Chem. in Britain* **9**, 490; 1973; and Thomas *Phil. Trans. R. Soc.* **A277**, 251; 1974). They and others have shown how various tactical substitutions in a molecular framework can be used as a steering device for preferred crystallization of a given desired crystal structure. The same approach has proved successful in persuading certain diolefins to crystallize with photoreactive rather than photostable structures, an approach which, in turn, has led to the solid-state synthesis of optically active molecules and many other novel products (Nakanishi *et al. Proc. R. Soc.* **A369**, 307; 1980).

Recently the delicate interdependence of crystal structure and reactivity in a family of related molecules derived from the skeletal framework shown below has been traced (Jones, Nakanishi, Theocharis and Thomas *J. chem. Soc. Chem. Commun.* 610; 1980). Various substitutions for X, Y and Z in the rings C, B and A were made and the photoreactivity of the monomer ascertained. The nature of the molecular packing was determined from the X-ray structure of each monomer crystal. It transpires that when the product (dimer) molecule occupies a similar volume within a crystal, and has a closely similar shape, as an adjacent pair of monomer molecules (an incipient dimer), single-crystal → single-crystal chemical transformation (without phase separation) can take place. In the figure below the adjacent molecules 1 and 2 (see framework with X, Y, Z all = H; this compound is 2-benzyl-5-benzylidene-cyclopentanone), which together constitute an incipient dimer, can undergo very minor movements at their peripheries during the course of a





Two neighbouring molecules of monomer (in which bonds are drawn with full lines) are so disposed with respect to one another in the 'engineered' crystal that they readily link up when irradiated with UV light to form a dimer. The four-membered cyclobutane ring (dotted bonds) is formed from the pairs of carbon atoms (shaded) linked originally by ethylenic bonds. Note how little the peripheral atoms (the benzene rings) move during this diffusionless reaction.

J. M. Thomas is Professor and Head of the Department of Physical Chemistry at the University of Cambridge.

0028-0836/81/080633-02$01.00

photodimerization that converts the doubly bonded carbon atoms (shaded) into a cyclobutane ring. In this crystal, monomer molecules pack so closely to one another — the ethylenic plane-to-plane distance is only 3.80Å — that dimerization proceeds freely to completion within the crystal matrix over a period of several hours.

Such changes are of considerable theoretical interest in that they permit, in principle, the atomic rearrangements to be monitored directly (at the diffusionless extreme) during the course of chemical reaction (Nakanishi, Jones, Thomas, Hursthouse and Motevalli; in the press). But such transformations have a wider relevance. New avenues of solid-state synthesis are opened up, by capitalizing on the idea that a potentially labile guest, when incarcerated within the well defined cell created by the crystal — be it of the pure material or of a host environment into which the guest has been buried on crystallization — has few options open to it. Through the agency of an externally applied stimulus, such as UV irradiation, the guest can to good advantage be converted into a more 'agreeable' state and then liberated (by dissolution for example) for appropriate utilization.

The sequential acts of molecular imprisonment and release, when prisoner and prison are carefully matched, can clearly be used to considerable advantage. Recent reviews (see Thomas *Pure appl. Chem.* **51**, 1065; 1979; Green, Lahav and Rabinovich *Acc. chem. Res.* **12**, 91; 1979) display the extent to which progress has been made in this direction. □

Habitat and the spread of rabies

from P.J. Bacon and D.W. Macdonald

AN international workshop on rabies, habitat classification and fox populations was held at Oxford recently*. As the behaviour of vectors of rabies, most notably the red fox, *Vulpes vulpes*, is known to be affected by the habitat and as the epizootiology of rabies appears to be affected by landscapes on both a fine and a gross level, it was hoped that a meeting of zoologists, epidemiologists and geographers could spark off new approaches to understanding and combatting a disease which has so far thwarted most attempts at control.

The first problem to be tackled was the classification of habitats. A robust method of describing landscapes is necessary not only for understanding the way foxes and rabies behave in differing habitats, but also for extrapolating from small study areas to the continental scale at which disease control must work. R.G.H. Bunce (Institute of Terrestrial Ecology, UK) described a classification method that was intended, in part, to enable the findings of detailed ecological studies at specific sites to be generalised to the whole of Britain. A wide variety of ecological parameters, such as soil profiles, natural vegetation and agricultural crops, determined by field-survey, are highly correlated to topographic features that can be measured from Ordnance Survey maps: these correlations can then be used to predict the factors from map measurements of areas not surveyed. The method accurately predicts, for example, acreages of various crops and potential primary productivity and an attempt is being made to produce a map of potential (that is, 'carrying-capacity') fox densities for the whole of Britain (Macdonald, D.W., Bunce, R.G.H. & Bacon, P.J. *J. Biogeography* in the press). The predictions are supported by the known fox densities in some areas although they cannot yet incorporate local variations in human hunting pressure. P.J.A. Howard, (Institute of Terrestrial Ecology, UK) reviewed methods of classification (that is categories concerned with natural divisions within the data) that were appropriate for habitat description; and concluded that while a suitable stratification into some type of category was desirable to permit efficient sampling of habitats, the aim of prediction might be better served by statistical methods using regression of the original data on variates important to the aspects being predicted. Thus, for example, it might be better to compare fox densities to measures of the landscape (for example, percentage of woodland, percentage of cereal crop) rather than to ordinal categories into which the landscape had been classified (for example, 'woodland', 'arable land' and so on). His recommendations received support from U. Emanuelsson (University of Lund) predictions of bird breeding densities in Swedish Lapland using formulae derived from regression of known densities against measures of landscape, such as area of woodland, altitude, and so on.

The aspirations of the workshop were well illustrated by quantitive investigations of the relations between landscapes and rabies spread. Using data on case-occurence made available by L. Andral (Centre d'etudes sur la Rage, France), F. Ball (University of Oxford), looked for relations between the velocity of the front wave of rabies to a set of data (collected by himself and P.J. Bacon) describing the habitats in which those cases occurred. Ball posed three questions (a) how does the behaviour of the front wave of rabies epizootic vary with habitat, (b) is infection better described by nearest neighbour models, that is with one resident fox passing the disease to its territorial neighbours and its own group members, or by dispersing fox models, that is, those taking into account the movements of many kilometres made by young male foxes, and (c) would a knowledge of habitat effects increase one's ability to predict the likely spread of the disease? He demonstrated that variations in front wave velocity can be linked to measures of the landscape, which correlate with fox densities. Examples of important measures are altitude (there are few foxes on mountain tops), railway lines (which follow fertile valleys and hence perhaps represent ribbons of high fox density, while also channeling fox movements) and vineyards (which are associated with broken hilly terrain suitable for foxes — indeed, biblical foxes often ate grapes and present day ones certainly eat blackberries). Map-data can thus be used to explain a significant amount of the historical variation of front-wave velocity in France, and similar measurements in areas in advance of the wave-front might have predictive value.

J. Ross (Imperial College London), applied B. Sayer's (Imperial College London), approach to the French data supplied by L. Andral and demonstrated the effects of rivers and motorways in diverting the path of the epizootic. Neither are readily crossed by foxes so movements tend to be along, not across them. From the case incidence data she drew trajectories of the disease front as it spreads across France month by month. When these trajectories are superimposed upon a geological map it is evident that the disease has followed a belt of limestone running south west across France. It seemed plausible that the rich soil of this limestone country was mirrored, through high primary productivity, in high fox populations. She also showed that rabies spread through areas of different landclasses (as described by Ball) had different temporal characteristics. These differences seemed largely associated with the presence or absence of upland pastures — habitat which might be good mouse-hunting country for foxes.

P.J. Bacon is at the Institute of Terrestrial Ecology, Merlewood Research Station, Cumbria and D.W. Macdonald is in the Department of Zoology, Oxford.

---

*The *Workshop in Rabies Spread, Habitat, Classification, and Fox Populations* was held at Keble College, Oxford from 15 to 19 September 1980. The sponsors were the Nuffield Foundation and the Insititute of Terrestrial Ecology. A full report of the meeting (with references and a list of participants) will be edited by the organisers and will be obtainable from The Librarian, Institute of Terrestrial Ecology, Merlewood Research Station, Grange over Sands, Cumbria LA11 6JU.

0028-0836/81/080634-02$01.00

Considerable debate revolved around the extent to which sophisticated analyses could be extended, considering the uncertainties of rabies surveillance. The reporting of cases is much more efficient along the epizootic front than in enzootic districts where familiarity soon produce negligence. Rather dismaying disclosures about the consistency of the data on which statisticians must work were endorsed in a devastating analysis by A. Wandeler (University of Berne) of the biases in the hunting figures frequently employed as indirect indicators of fox numbers. Factors such as the phase of the moon affect hunters' behaviour, and hence the bag records! It was widely agreed that every effort should be made to standardise the recording of rabies incidence, even if only in sample districts.

Students of fox ecology were not the only biologists present, since the principles underlying analyses of habitat utilisation are widely applicable. H. Kruuk, (Institute of Terrestrial Ecology, UK), showed how the dispersion of habitat patches (and hence available food in the form of earthworms) determined the size of badger *(Meles meles)* territories and their local population density. Earthworms can only be captured on warm and humid nights when they are tempted onto the surface. From night to night, with shifting wind direction for example, different fields will temporarily become bountiful foraging sites. Kruuk argued that badger territories must contain sufficient foraging sites so that at least one will yield worms irrespective of changes in the wind. Territory size, and ultimately population density, is thus a function of the dispersion of potential earthworm patches — a factor that can to some extent be predicted from maps. When the meeting learnt from Wanderler (University of Berne) that badgers had been more involved than previously thought in rabies spread in Switzerland it was realised that these findings were of practical as well as theoretical importance. Kruuk's findings in Scotland could be compared with those of I. Lindsay and D. Macdonald (University of Oxford). They had examined the habitat components of badger territories in Buckinghamshire and found, for example, that territory size varied with the acreages of cereal land they encompassed. These arable crops are unimportant in the badgers' diet and can be interpreted as 'dead space' in the badgers' territories. S. Harris (University of Bristol), found urban-dwelling badgers living at extraordinarily high densities. Indeed, the participants collectively blanched at the prospect of rabies control in urban areas, because it was apparent that the densities of both wild and domestic vectors are extremely high and almost unassailable by conventional control methods. The highest fox densities recorded were from suburbs where Macdonald has found home ranges as small as 10 hectares. In the same habitat foxes were living in social groups — another factor which doubtless would effect rabies epizootiology.

Two schools of thought emerged on the most useful way to proceed — one was that fox population dynamics were so complicated that effort might most profitably be directed to seeking broad relationships between general features of the landscape and rabies spread, while the opposing view was that the data on rabies spread were so coarse and inconsistent that direct and detailed study of foxes was more likely to yield a better understanding of the factors affecting the pattern of rabies spread. In either case, concensus was reached that the most crippling obstacle was our ignorance of the behaviour of rabid foxes and the extent to which it differs from that of healthy animals. □

NATURE

100 years ago

The Singular Methods of Travel the Wagtail adopts to Cross the Mediterranean Sea.

One evening we were sitting at the foot of the pyramid of Cheops, sipping our cup of fragrant Mocha and in jolly conversation, rolling up clouds of blue smoke from our Korani cigarettes. We were waiting for the sinking of the sun to make our return to Cairo. The deep silence of the surrounding desert possessed something uncommonly solemn, only now and then disturbed by the cry of the hoarse fishhawks far above us. Still higher the pelicans were grandly circling. Their flight, though heavy when seen from anear, possess a majesty in the distance attained by no other bird. Right before us several wagtails were hopping around and 'tilting'. They were quite tame, and flew restlessly hither and thither. On this occasion I remarked, 'I could not quite understand how these birds could make the long passage of the Mediterranean.' Sheik Ibrahim heard this from our interpreter. The old Bedouin turned to me with a mixture of French and Arabic as follows, which the interpreter aided us to fully comprehend:-

" 'Do you not know, Hadretch (noble sir), that these small birds are borne over the sea by the larger ones?'

"I laughed, as did our friends; for at first we thought we had misunderstood him; but no: the old man continued quite naturally:-

" 'Every child among us knows that. These little birds are much too weak to make the long sea journey with their own strength. This they know very well, and therefore wait for the storks and cranes and other large birds, and settle themselves upon their backs. In this way they allow themselves to be borne over the sea. The large birds submit to it willingly; for they like their little guests, who by their merry twitterings help to kill the time on the long voyage.'

"It appeared incredible to us. We called to a pair of brown Bedouin boys, pointed out the wagtails to them, and inquired:-

" 'Do you know whence come these small birds?'

" 'Certainly,' they answered. 'The Abu Saad (the stork) carried them over the sea.'

"At supper, in the Hôtel du Nil, I related the curious story to all present, but naturally enough found only unbelieving ears.

"The only one who did not laugh was the Privy Councillor Heuglin, the famous African traveller, and, excepting Brehm, the most celebrated ornithologist of our time for the birds of Africa. I turned to him after the meal, and inquired of his faith. The good royal councillor smiled in his caustic way, and with a merry twinkle remarked: 'Let the others laugh: they know nothing about it. I do not laugh, for the thing is known to me. I should have recently made mention of it in my work if I had had any strong personal proof to justify it. We must be much more careful in such things than a mere story-teller or novel-writer; we must have a proof for everything. I consider the case probable, but as yet cannot give any warrant for it.'

"My discovery, if I may so call it, I had kept to myself, even after Heuglin had thus expressed himself, and would even now maintain silence on the subject had *I not* recently discovered a new authority for it."

I read lately in the second edition of Petermann's great book of travels the following:-

"Prof. Roth of Munich related to me in Jerusalem that the well-known Swedish traveller, Hedenborg, made the following interesting observation on the Island of Rhodes, where he stopped. In the autumn tide, when the storks come in flocks over the sea to discover them. Once he followed a flock of storks, and as they lighted he saw small birds fly up from their backs, which in this manner had been borne over the sea. The distance prevented him from observing to which species of singing birds they belonged."

We regret that the Lords should have thrown out the Bill on Tuesday for the Opening of Museums and similar places on Sundays. As the *Times* very well puts it:— "The gravity of the question is that London has in its midst people to whom anything of the nature of intellectual toil — and prolonged sight-seeing is of that character — is essentially irksome. But they are human beings, and not lost to all salutary influence. It would be folly to despair of making the Sunday more tolerable than it is to them. Our climate does not often admit of men and women sitting out of doors talking or listening to elevating music. Some substitute must be found to put on on equality with the people of more sunny lands. It is the task of true friends of the working classes to suggest means by which, without any revolution in national ideas as to the sacredness of Sunday, they may be enabled to taste those simple and primitive pleasures — for example, the pleasure of pure repose of mind and body, or that of hearing music — which all, even the untutored, can enjoy.

From *Nature* 23, 24 February, 387, 1880.

0028-0836/81/080635-01$01.00

The expanding science of photobiology

from John Jagger

Photobiology is concerned with the vast range of actions of nonionizing radiations on biological systems — from photolysis of simple molecules in pure solution to the effects of sunlight on human skin and problems of world food production through photosynthesis. The common thread that held the five hundred participants at the recent International Photobiology Congress at the Louis Pasteur University, Strasbourg, together is an interest in the fundamental biophysical mechanisms involved in the interactions of light and living matter. In line with this orientation is the increasing use of sophisticated physical techniques, such as picosecond dye-laser pulse studies of electron energy exchange between chlorophyll pigments in chlorophyll–protein complexes (J. Breton and G. Paillotin, Centre d'Etudes Nucléaires de Saclay; N.E. Geacintov, New York University), and picosecond kinetic measurements at liquid-helium temperature of proton translocation in the retinal molecule (M. Chabre, Grenoble).

Reversible photochemical reactions are often used by biological systems to sense changes in the environment. The most obvious and certainly the most sophisticated such system involves the protein rhodopsin pigments used in **animal vision**. Photoisomerization of the retinal prosthetic group of rhodopsin leads to an alteration of the receptor cell membrane potential, but the intermediate steps are not understood. Since the cell membrane is not connected to the disks that contain the rhodopsin, there must be a diffusible transmitter, and this is now believed to be cyclic GMP, rather than the previously proposed calcium ion (M. Chabre, Grenoble). It is proposed that rhodopsin floats in a fluid membrane and, on photoexcitation, interacts with an enzyme chain controlling cyclic GMP levels. There is evidence that the site of interaction of rhodopsin with membrane enzymes is remote from the retinal itself, which implies that there is an overall conformational change of the whole rhodopsin molecule. Support for this picture comes from the findings of H. Shichi and R.L. Somers (National Eye Institute) that purified rhodopsin illuminated *in vitro* activates a purified GTPase; this results from formation of a protein–rhodopsin complex and does not require a membrane environment.

Another important reversible photoreaction used by biological systems for sensory purposes is the **phytochrome system** of plants. H. Smith (University of Leicester) discussed the role of phytochrome in the natural environment. Light penetrating a plant canopy is enriched in far-red wavelengths (700–750 nm), because of absorption of other wavelengths by chlorophyll. This alters the phytochrome photoequilibrium, thus permitting the plant to detect the amount of canopy shade, and to react accordingly — either by growth out of the shaded area or by changing leaf and chloroplast development and metabolism to adapt to the shade regime. However, it is still not clear whether this detection depends on the ratio of the red and far-red forms of phytochrome or on the amount of the far-red form present.

Extremely sensitive **bioluminescence assays** were described by A.M. Michelson (Institut de Biologie Physico-Chimique, Paris), based upon the fact that clam luciferases behave like peroxidases, producing superoxide radical anion, which then causes the clam luciferin to luminesce. Such systems permit estimation of femtogramme quantities of peroxide, and exquisitely sensitive estimations of hydrogen peroxide, superoxide radical anion, and enzymes like peroxidase and superoxide dismutase. In enzyme assays, turnover of the enzyme provides an effective amplification of the light. Such techniques have also been used for estimation of antibodies bound either to such enzymes or to luciferases.

In the area of basic mechanisms in **ultraviolet photobiology**, progress has been made in the study of lethal effects of near-UV radiation (300–380 nm), which is present in sunlight at the surface of the Earth. M.J. Peak and J.G. Peak (Argonne National Laboratory) have shown that there is a 365 nm maximum in the action spectrum for induction of single-strand breaks in purified bacterial transforming DNA irradiated *in vitro*. These breaks are primarily true breaks (as opposed to alkali-labile bonds) and generally require oxygen; they might explain the small plateau observed at this wavelength in action spectra for killing of *E.coli* in air. However, work by A.G. Miguel and R.M. Tyrrell (Federal University of Rio de Janeiro) indicates that single-strand breaks are very readily repaired and are therefore probably not the major lethal lesion. Similar arguments have previously eliminated pyrimidine dimers from consideration. Thus, the primary near-UV lethal lesion in wild-type bacteria remains to be identified. Whatever the damage, one of its first manifestations appears (S.H. Moss and K.C. Smith, Stanford University School of Medicine) to be an altered cell membrane, which makes a significant contribution to the killing of cells plated on minimal medium.

An inducible repair system for near-UV lethal damage in *E. coli* has been found by J. Peters and J. Jagger (University of Texas at Dallas). The system accounts for a large fraction of lethality in actively growing cells, and is distinct from the far-UV-induced SOS system. In related work, S. Menezes and R.M. Tyrrell have shown that repair systems inducible by far-UV (200–300 nm) radiation play a relatively minor role in the repair of near-UV damage.

Photomedicine was much in evidence reflecting a maturation of photobiology in its application to medical problems. Of great interest were developments in so-called PUVA therapy, which involves the treatment of skin lesions, such as psoriasis, by administering the dye *P*soralen (orally) and then irradiating with *UV-A* light (near-UV light in the range 320–380 nm). This therapy has been extremely effective in producing clearance and remissions. Unfortunately, there is concern that long term PUVA may act as a promoter of non-melanoma skin cancer, although this hazard has not prevented some 80,000 patients in all parts of the world from utilizing the therapy because the disease itself is so disfiguring. P. Vigny, D. Averbeck and co-workers (Institut Curie, Paris) have now found that a psoralen derivative, 3-carbethoxypsoralen, does not produce interstrand DNA cross-links, in contrast to the commonly used 8-methoxypsoralen, and the new compound shows therapeutic activity in mice without exhibiting carcinogenesis. Additional new isopsoralen derivatives that produce only monofunctional addicts in DNA have been synthesized by the Padua University group of G. Rodighiero. This is a happy development which opens the way to possible broad use of light-and-chemical treatments (photochemotherapy) of skin disorders.

A second area of great medical interest concerns the immunological rejection of UV-induced **skin cancers**, studied by M.L. Kripke and co-workers (Frederick Cancer Research Center, Maryland). Many UV-induced squamous carcinomas and fibrosarcomas are immunologically rejected when transplanted to normal syngeneic recipients, but grow progressively in immunologically deficient animals or animals that have been UV irradiated (including the animals in which they were originally produced). Kripke and co-workers have found that antigens produced during UV irradiation of the skin trigger the production of suppressor T lymphocytes, which in turn prevent an immunological rejection of the tumour. Langerhans cells (macrophage-like cells in the skin) may be involved in the initial events in this process. K. Wolff (Innsbruck University) pointed out that Langerhans cells show the highest susceptibility to UV damage of all skin cells, and that UV

John Jagger is a Professor in the University of Texas at Dallas.

0028-0836/81/080636-02$01.00

abolishes their immunological surface markers and abrogates their capacity to stimulate T-lymphocyte proliferation. E.C. DeFabo and F.P. Noonan (Kripke's group) have found that the wavelength maximum for depression of contact sensitivity in mouse skin is at about 270 nm, which may reflect absorption by urocanic acid, a metabolite of histidine, which accumulates in the skin where it cannot be further metabolized. Urocanic acid, incidentally, has been proposed by J.A. Parrish (Massachusetts General Hospital) to act as a natural photoprotection (light-screening) agent, which is partially removed when skin is washed, explaining an observation first made by Jacopo Beccari in 1744 that his hands fluoresced better in sunlight after being washed.

Finally, a disturbing medical development is the recent rise in the incidence of malignant melanoma: world mortality rates have approximately doubled in the last fifteen years. It is not generally recognized that skin cancers are the most common malignant tumours of man. Although the overall mortality is less than 5%, malignant melanomas show about a 40% mortality, and they represented about 1% of all cancer deaths in 1970. The recent increase in melanoma incidence is generally attributed to increased sun-exposure habits, although it is possible that there is some kind of synergism between UV radiation and environmental toxic chemicals, as in the observations of P.D. Forbes and co-workers (Temple University Health Sciences Center) that concurrent exposure to retinoic acid and UV stimulates tumour development. Effective topical sunscreens are thus of great importance, and the need for international standardization of sunscreen efficacy was discussed in a Round-Table held at the congress.

Photobiology has become a very active discipline that is now rapidly maturing. Sophisticated physical and chemical techniques are being applied, and significant advances in fundamental knowledge are being made. It is playing an increasingly important role in medicine. All this comes as no surprise to the photobiologists, who share the view of Isaac Asimov that "The Earth is bathed in light from the Sun, and one could scarcely think of a more important single fact than that." □

[illegible] Upper Circular Road [illegible]

Antisickling drugs and water structure

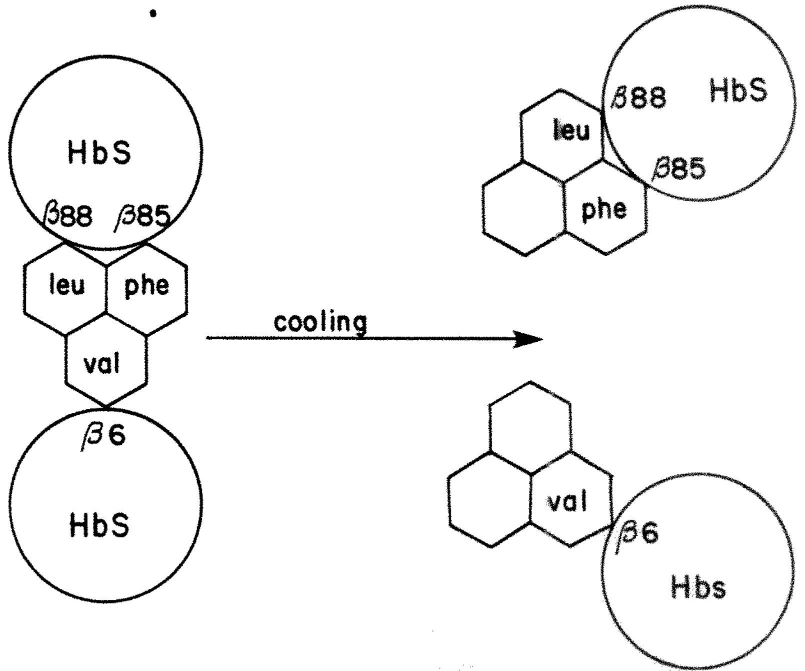
from Ruth and Reinhold Benesch

In sickle cell anaemia the substitution of a valine for a glutamic acid in the β chain of the normal haemoglobin molecule causes a polymerization of the deoxygenated sickle cell haemoglobin. In order to design drugs to prevent aggregation of the sickle cell haemoglobin (HbS) an understanding of the molecular mechanisms involved is obviously required[1,2]. Stillinger's[3] recent article, *Water Revisited*, has led us to suspect that the hydrophobic class of sickling inhibitors may act on water structure rather than by stereospecific interaction with HbS itself. Stillinger points out the analogy between hydrophobic interaction and the anomalous behaviour of supercooled water. Water is described as a random three-dimensional network of hydrogen bonds in which are embedded relatively bulky, unstrained hydrogen bond polyhedra. The concentration of these polyhedra increases as the temperature is lowered. The attractive force between them and their consequent 'clumping' is responsible for some of the unusual properties of supercooled water. When a non-polar solute is dissolved in water, the random network of hydrogen-bonded water molecules reorganizes around it to form very similar bulky convex polyhedra which only differ from the empty ones found in supercooled water by the presence of the hydrophobic molecule in their centre. The attraction between these 'filled' convex polyhedra provides the physical basis for the 'hydrophobic bond'.

A characteristic feature of hydrophobic interactions is their unusual temperature dependence, that is, they become weaker as the temperature is lowered. This can readily be explained in terms of the picture provided by Stillinger. Since lowering the temperature will increase the concentration of empty polyhedra, the probability of a filled one interacting with an empty one, rather than with another filled one, will increase and the attraction between the hydrophobic residues will weaken with decreasing temperature.

When sickle cell haemoglobin polymerizes it forms a gel that 'melts' as the temperature is lowered. Crucial side-to-side contact between strands of HbS molecules in the crystal (and therefore presumably also in the fibres formed by the haemoglobin) was shown to exist between valine β6 of one molecule and phenylalanine β85 and leucine β88 of the neighbouring one[4]. If this bond can now be regarded as the result of the attractive forces between hydrogen bond polyhedra containing these residues (see the figure), then it should be broken as the temperature is lowered because of competition by the empty water polyhedra as their concentration increases with decreasing temperature.

An analogous mechanism would also account for the well documented inhibition of HbS polymerization by a variety of compounds containing hydrophobic residues[5-9]. There is no clearcut relationship between the structure of these inhibitors and the known geometry of the sites containing valine β6 and phenylalanine β85



Depolymerization of deoxy-HbS on cooling: In this schematic diagram the hexagons represent the hydrogen bond polyhedra discussed in the text.

Ruth and Reinhold Benesch are in the Department of Biochemistry, Columbia University, New York.

0028-0836/81/080637-02$01.00

and leucine β88. Thus, for example, the D isomeric forms of aromatic amino acids are as active as the L isomers[10]. This supports the hypothesis that the inhibition of polymerization by these compounds is brought about by competition between the water polyhedra filled with the small hydrophobic molecules with those attached to the protein, rather than a direct interaction between the hydrophobic residues and the protein which should be stereospecific. This picture would also explain why attempts to increase the effectiveness of hydrophobic antisickling agents by structural modification have met with relatively little success. □

1. Schechter, A. N. *Hemoglobin* **4**, 335 (1980).
2. *Development of Therapeutic Agents for Sickle Cell Disease INSERM Symp.* (eds Beuzard, Y., Hercules, J. & Rosa, J.) (Elsevier, Amsterdam, 1979).
3. Stillinger, F. H. *Science* **209**, 451 (1980).
4. Wishner, B. C., Ward, K. B., Lattman, E. E. & Love, W. E. *J. molec. Biol.* **98**, 179 (1975).
5. Elbaum, D., Nagel, R. L., Bookchin, R. M. & Herskovits, T. T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4718 (1974).
6. Ross, P. D. & Subramanian, S. *Biochem. biophys. Res. Commun.* **77**, 1217 (1977).
7. Votano, J. R., Gorecki, M. & Rich, A. *Science* **196**, 1216 (1977).
8. Kubota, S. & Yang, J. T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5431 (1977).
9. Behe, M. J. & Englander, S. W. *Biochemistry* **18**, 4196 (1979).
10. Noguchi, C. T. & Schechter, A. N. *Biochemistry* **17**, 5455 (1978).

Fifty years of Slough's ionosphere

from Henry Rishbeth

ON 11 January 1981 the 'noon run' of the Slough ionosonde marked with due ceremony the fiftieth anniversary of routine ionospheric sounding. An ionosonde is a kind of swept-frequency radar and radar developed from the experimental studies of the ionosphere that were conducted at Slough and elsewhere in the twenties and thirties. This work followed from the basic experiments of 1924–25 in Britain (E. V. Appleton and M. A. F. Barnett *Nature* **115**, 333; 1925) and in the US (G. Breit and M. A. Tuve *Nature* **116**, 357; 1925) that conclusively proved the existence of radio-reflecting layers in the upper atmosphere. It was Watson-Watt at the Radio Research Station (RRS), Slough (now the SRC Appleton Laboratory) who coined in 1926 the term 'ionosphere' (G. W. Gardiner *Nature* **224**, 1096; 1969). Years later, the output of analysis and interpretation based on Slough data gave rise to the unofficial definition of the ionosphere as 'the part of the atmosphere vertically above Slough'.

To promote scientific understanding of the ionosphere and to meet the growing needs of radio communications, systematic data were needed. RRS publications contain some h'–f data on apparent reflection heights (h') at different radio frequencies (f) obtained in 1930. However, the most important data on the ionospheric layers (E, F1 and F2 layers, respectively at heights of around 100, 180 and 250–350 km) are their 'penetration' or 'critical' frequencies (fE, fF1, fF2) as measured by a swept-frequency sounder. The series of critical frequency data began at noon on 11 January 1931 (a Sunday, like its fiftieth anniversary) and has continued with few breaks to the present day. Initially only fE was recorded, on ninety days in 1931. Tabulated fF2 data start on 19 December 1931, but it was not until 20 September 1932 that the long series of daily values was initiated, the first 'noon run' for which all three parameters fE, fF1, fF2 are preserved being on 23 September 1932. Hourly soundings — which are now made routinely — were made only for three 24-hour periods during 1931; from these data it appears that the initial value at 1200 GMT on 11 January 1931, namely 2.1 MHz, may be about 15–20 per cent low in comparison with neighbouring values. Furthermore, it was then also possible to record fE at night to values as low as 0.7 MHz: this could not be done nowadays with normal ionosondes in England, owing to the prevalence of radio broadcast interference (R. W. Smith, personal communication). Today, data from the Slough ionosonde are distributed to numerous users on schedules to meet their particular needs, which may include long-term 'predictions' of ionospheric conditions or short-term 'forecasts' of disturbances. Some users, including the BBC, Marconi's and Jodrell Bank, receive noon 'critical frequency' data by telex each working day. It is interesting that the largest value of fF2 ever recorded at Slough occurred in the present solar cycle, namely 17.2 MHz at 1300 and 1400 GMT on 15 December 1979.

Other long series of ionospheric data have been produced at Huancayo (Peru), Watheroo (Western Australia) and Washington (DC). By the start of the International Geophysical Year of 1957–58 about 140 ionospheric observatories were in operation, many of which have continued to this day, thereby creating global data sequences extending over two sunspot cycles. This well organized scientific resource is available to researchers through the World Data Centre at Slough, and corresponding Centres in Boulder (Colorado), Moscow and Tokyo, established at the start of the IGY.

What of the future? Following the merger with the SRC Rutherford Laboratory, Appleton Laboratory work is being transferred away from Slough. Current thinking is that an ionosonde should remain permanently in the Slough area. Possibilities being considered for the future are the automatic read-out of data to the merged laboratories' site at Chilton, with dial-up access by users both to past data and to the up-to-the-hour state of 'the part of the atmosphere vertically above Slough'.

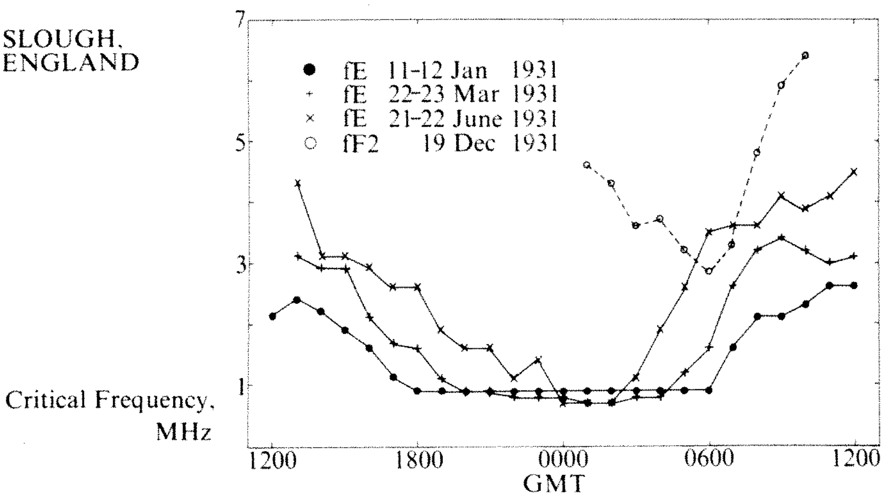


Fig. The four sequences of hourly critical frequency data recorded at Slough (51½°N, ½°W) during 1931
The values refer to the 'ordinary' magnetoionic component. Some of the fluctuations in fE are probably due to sporadic E.

Henry Rishbeth is at the Rutherford and Appleton Laboratories, Slough, UK.

0028-0836/81/080638-01$01.00

ARTICLES

Stable isotope composition of benthic foraminifera from the equatorial Pacific

Edith Vincent, J. S. Killingley & W. H. Berger

Scripps Institution of Oceanography, La Jolla, California 92093

When ranked according to deviation from isotopic equilibrium, deep-sea benthic foraminifera exhibit no correlation between the rankings for oxygen and for carbon. Both rank sequences are stable from east to west in the equatorial Pacific (fertility-independent) and from Glacial to post-Glacial (time-independent). Records for the past 15,000 yr suggest marked changes in deep-sea circulation during this time interval.

STABLE isotopic compositions of benthic foraminifera reflect bottom water conditions. The $^{18}O/^{16}O$ ratio of the calcium carbonate in their shells is strongly dependent on the temperature and oxygen isotope composition of the ambient water, whereas the $^{13}C/^{12}C$ ratio is related to the carbon isotopic composition of the total dissolved inorganic carbon. Oxygen isotopic changes in benthic foraminifera from deep-sea sediments, first used in Emiliani[1,2] to reconstruct bottom water changes, have been used to reconstruct Pleistocene variations in Atlantic bottom water temperatures[3,4]. But changes in the carbon isotopic composition of benthic foraminiferal tests can also aid palaeoceanographic interpretations[5].

Although the initial assumption that all deep-sea benthic foraminifera deposit their shells in equilibrium with seawater[6] is incorrect[7-10], it is thought that for oxygen isotopes, deviation from equilibrium is small in some species[11,12], and that more substantial deviations of other species are constant and can be corrected for by adding a constant value to the measured $^{18}O/^{16}O$ ratio[5,13-19]. We show here that the ranges of variation within individual species and of isotopic differences between species are of the same order of magnitude as isotopic changes commonly recorded in Pleistocene sediments. Thus, our results impose limits on the significance of downcore isotope records based on single-species analyses, especially if small samples with few specimens are used.

Materials and methods

Benthic foraminifera from western and eastern equatorial Pacific sediments, recovered by box coring during Scripps Institution of Oceanography Expeditions Eurydice (ERDC, 1975) and Pleiades (PLDS, 1976) respectively, were analysed. Sedimentation patterns in the two areas have been described elsewhere[20,21]. They are characterized by relatively high sedimentation rates and high carbonate content (Table 1). Carbonate content in the east (PLDS) is lower than in the west (ERDC), largely because of high silica content in the eastern high fertility area. The higher fertility in the east is also reflected by the rate of sedimentation and a high ratio of *Neogloboquadrina dutertrei* to *Pulleniatina obliquiloculata* (Table 1), the former planktonic species being associated with upwelling[22]. For the present study, dissolution effects are not important above 3,000 m depth, and probably did not affect our results significantly, even in the two deepest cores.

We sampled the surface sediments in these box cores as well as a deeper level in Box-core ERDC-92 dated at 16,000–17,000 ^{14}C yr (ref. 23). All benthic foraminifera larger than 250 μm in each sample were selected and 150–400 specimens per sample were thus isolated. Of these specimens, 10 species were abundant enough to run replicate isotopic analyses (Table 2). Over 100 analyses for both oxygen and carbon were performed. Typically 4–7 specimens with a total weight of at least 0.5 mg were used for each analysis. Samples were cleaned ultrasonically to remove fine particles from inside the foraminiferal tests. Isotopic analysis followed standard procedures[24]. Benthic foraminifera taxonomy followed that of Culp[25].

In Core ERDC-92 foraminifera were analysed from discrete size fractions (149–250 μm, 250–420 μm and >420 μm). No correlation of isotopic composition with size was detected, however, in contrast to the well-established size effect on isotopic composition in planktonic foraminifera[26].

Species abundances

The two most common benthic species in samples shallower than 3,000 m are *Planulina wuellerstorfi* and *Oridorsalis umbonatus*, which each comprise 10–20% of the total number of individuals in the calcareous benthic assemblage. Below 3,000 m, these species and *Nuttalides umbonifera* dominate the benthic assemblages (ERDC-135 and PLDS-66). Here *N. umbonifera* represents 12–13% of the calcareous benthic fauna, whereas it is absent in the shallower samples.

All other species analysed are less common and generally constitute <5% of the calcareous benthic fauna. *Uvigerina*, which is rare (<2%) or absent in all ERDC samples, is common in the PLDS-66 sample, making up 15% of the calcareous benthic fauna. The higher productivity of surface waters in the eastern region is presumed to cause an increased supply of food to the underlying benthos. Therefore we suggest that the abundance of *Uvigerina* indicates a region of high fertility. This

Table 1 Core location and sample material

Box core no.	Lat	Long	Water depth (m)	Depth in core (cm)	Age	Sedimentation rate* (cm kyr^{-1})	$CaCO_3$	dut./Pul.†
ERDC-92	2°13.5′S	156°59.9′E	1,598	0–1.5	Holocene	1.6	90%	<1
ERDC-92	2°13.5′S	156°59.9′E	1,598	20–30	Glacial	1.8	86%	5
ERDC-123	0°01.3′S	160°24.9′E	2,948	0–1	Holocene	1.8	84%	<1
ERDC-135	0°52.5′N	160°59.6′E	3,509	0–2	Holocene	1.4	85%	<1
PLDS-66	0°56.6′N	104°6.1′W	3,496	0–2	Holocene	2.0	72%	72

* Derived from ^{14}C dates (ERDC-92, ref. 23; ERDC-123, 135, ref. 32; PLDS-66, unpublished data).

† Ratio of *N. dutertrei* to *P. obliquiloculata* (>149 μm).

0028-0836/81/080639—05$01.00

Table 2 Isotopic composition of various benthic foraminiferal species

Groups	Species	(1) ERDC-92, Holocene		(2) ERDC-92, Glacial	(3) ERDC-123, Holocene		(4) ERDC-135, Holocene		(5) PLDS-66, Holocene	
a $\delta^{13}O$										
"Low ^{18}O"	*Nuttalides umbonifera*	—		—	—		3.00 (3) *r* 0.84	−0.60	2.43 (3) *r* 0.30	−1.17
	Laticarinina pauperata	2.65 (6) *r* 0.84 ± 0.29	−0.55	3.10 (2) *r* 0.67	2.89 (1)	−0.61	3.20 (1)	−0.40	2.69 (3) *r* 0.46	−0.91
	Planulina wuellerstorfi	2.69 (12) *r* 1.12 ± 0.42	−0.51	3.13 (5) *r* 1.14 ± 0.51	2.66 (3) *r* 0.21	−0.84	3.21 (2) *r* 0.45	−0.39	2.50 (4) *r* 0.06 ± 0.04	−1.10
	Favocassidulina favus	—		—	2.98 (1)	−0.52	3.09 (3) *r* 0.46	−0.51	—	
"High ^{18}O"	*Oridorsalis umbonatus*	3.23 (8) *r* 1.58 ± 0.49	+0.03	3.99 (3) *r* 0.49	3.34 (3) *r* 0.22	−0.16	3.66 (3) *r* 0.65	+0.06	3.33 (3) *r* 0.46	−0.27
	Siphouvigerina interrupta	3.25 (2) *r* 0.36	+0.05	3.99 (2) *r* 0.44	—		—		—	
	Cassidulina subglobosa	3.33 (4) *r* 1.01 ± 0.43	+0.13	4.55 (4) *r* 0.27 ± 0.12	3.33 (1)	−0.17	3.70 (1)	+0.10	3.18 (3) *r* 0.31	−0.41
	Quinqueloculina auberiana	3.21 (2) *r* 0.07	+0.01	4.30 (2) *r* 0.07	—		—		—	
	Uvigerina spp.	—		—	—		—		3.16 (4) *r* 0.64 ± 0.27	−0.44
	Hoeglundina elegans	4.04 (5) *r* 0.67 ± 0.27		4.79 (4) *r* 0.70 ± 0.29						
Estimated equilibrium*		3.2		—	3.5		3.6		3.6	
b $\delta^{13}C$										
"Super low ^{13}C"	*Oridorsalis umbonatus*	−0.53 (8) *r* 0.53 ± 0.18	−1.93	−0.75 (3) *r* 0.79	−0.89 (3) *r* 0.36	−2.29	−0.93 (3) *r* 0.29	−2.43	−1.17 (3) *r* 0.35	−2.37
	Favocassidulina favus	—		—	−0.66 (1)	−2.06	−0.61 (3) *r* 0.23	−2.11	—	
"*Low* ^{13}C"	*Uvigerina* spp.	—		—	—		—		−0.63 (4) *r* 0.48 ± 0.21	−1.83
	Siphouvigerina interrupta	−0.14 (2) *r* 0.22	−1.54	−0.47 (2) *r* 0.07	—		—		—	
	Cassidulina subglobosa	0.04 (3) *r* 0.37	−1.36	−0.49 (4) *r* 0.47 ± 0.22	−0.59 (1)	−1.99	−0.56 (1)	−2.06	−0.61 (3) *r* 0.36	−1.81
"Medium ^{13}C"	*Laticarinina pauperata*	−0.1 (6) *r* 1.41 ± 0.48	−1.41	−0.20 (2) *r* 0.25	−0.18 (1)	−1.58	−0.28 (1)	−1.78	−0.21 (3) *r* 0.09	−1.41
	Nuttalides umbonifera	—		—	—		−0.23 (3) *r* 0.70	−1.73	−0.18 (3) *r* 0.33	−1.38
"High ^{13}C"	*Planulina wuellerstorfi*	0.40 (12) *r* 0.49 ± 0.12	−1.00	0.45 (5) *r* 0.33 ± 0.15	0.24 (3) *r* 0.28	−1.16	0.07 (2) *r* 0.02	−1.43	0.19 (4) *r* 0.39 ± 0.18	−1.01
	Quinqueloculina auberiana	0.64 (2) *r* 0.25	−0.76	1.13 (2) *r* 0.16	—		—		—	
	Hoeglundina elegans	2.13 (5) *r* 0.29 ± 0.11		2.14 (4) *r* 0.32 ± 0.16	—		—		—	
Estimated equilibrium†		1.4		—	1.4		1.5		1.2	

All specimens are >250 μm except *S. interrupta*, which is from the 149–250 μm size fraction. In each column the first number to the left is the mean δ value with respect to ‰ PDB; the number in parentheses is the number of analyses followed by the range *r* and by the standard deviation if more than three analyses were made; the last number, ro the right, is the Δδ value (the mean δ value minus the estimated equilibrium value).

* For present-day conditions the $^{18}O/^{16}O$ composition of calcite at equilibrium with ambient-bottom water was calculated using a palaeotemperature equation[44] and estimated temperatures[45] of 3°, 1.8°, 1.5° and 1.5 °C for ERDC 92, 123, 135 and PLDS 66, respectively. The isotopic composition of the bottom water was assumed to be −0.4 (PDB) (ref. 46).

† Present-day equilibrium $^{13}C/^{12}C$ composition of equilibrium calcite was estimated for the four core sites using reported fractionation relationships between calcium carbonate and dissolved bicarbonate[47]. Ambient $\delta^{13}C$ (PDB) values for bicarbonate were calculated as 0.15, 0.18, 0.25, and 0.25 for cores ERDC 92, 123, 135 and PLDS 66, respectively, from apparent oxygen utilization (AOU) to $\delta^{13}C$ relationships[13]. AOU values were obtained from reported temperature, salinity and dissolved O_2 data[45,48].

assessment is supported by its decreasing abundance in Box-Core ERDC-123 from Glacial to Holocene sediments (from 6% of the benthic calcareous assemblage to <1%) in parallel to other trends commonly associated with decreasing fertility: (1) a decreasing sedimentation rate (from 3 to 1.8 cm kyr^{-1}); (2) a decreasing *N. dutertrei*/*P. obliquiloculata* ratio (from >1 to 0.2); and (3) a decreasing benthic foram number (from ~30 to 10 benthic foraminifera larger than 250 μm g^{-1} of dry sediment).

Oxygen isotopes

Results of the oxygen analyses are given in Table 2*a*. Species are listed in the order of deviation from estimated equilibrium values, which for $\delta^{18}O$ in present day conditions range from 3.2 to 3.6‰, depending on the water depth of the core location (see Table 2*a*). Our comparisons are between species, between periods (ERDC-92, Holocene and Glacial), and between areas of low fertility in the west (ERDC) and high fertility in the east (PLDS). The disequilibrium ranking (greatest disequilibrium first) consists of two groups: 'low ^{18}O' and 'high ^{18}O'. The low ^{18}O group includes *N. umbonifera*, *L. pauperata*, *P. wuellerstorifi*, and *Favocassidulina favus*, which are clearly depleted in ^{18}O relative to equilibrium by around 0.5‰ in the ERDC samples. The high ^{18}O group contains *O. umbonatus*, *Siphouvigerina interrupta*, *Cassidulina subglobosa*, *Quinqueloculina auberiana* and *Uvigerina* spp., which have oxygen values close to equilibrium (for *Uvigerina* see ref. 5). *Hoeglundina elegans* is a separate case: its shell is aragonitic[27–29]. The equilibrium values for aragonite are not known for these low temperatures.

The grouping into low ^{18}O and high ^{18}O seems unrelated to time (column 1 compared with column 2) and with fertility (columns 1, 3, 4 compared with column 5). From the $\Delta\delta^{18}O$ values (the mean δ values minus the estimated equilibrium values) in Table 2*a* and Fig. 1*a*, it seems that there is a fertility effect on isotopic composition, because the degree of disequilibrium for oxygen increases from west to east. Compared with the eastern core (PLDS 66), the five benthic species shown for the western cores (ERDC) all contain a higher $^{18}O/^{16}O$ ratio and are closer to estimated equilibrium values. Moreover, between the western cores an increase in disequilibrium with increasing sedimentation rate is indicated, which in turn is related to fertility.

The large range in most of the replicate analyses makes it impossible to confirm that the amount of change in disequilibrium with fertility is species dependent.

Carbon isotopes

Results of the carbon analyses are given in Table 2*b*. Again, species are listed in the order of deviation from estimated equilibrium values, which range in $\delta^{13}C$ from 1.2 to 1.5‰, depending on the apparent oxygen utilization (AOU) in the water. The AOU indicates how much oxygen has been used to oxidize organic matter to produce additional CO_2 (with a $\delta^{13}C$ near −20‰)[26,30]. The ranking (greatest disequilibrium first) consists of four groups: 'super low ^{13}C', 'low ^{13}C', 'medium ^{13}C', and 'high ^{13}C'. The super low ^{13}C group includes *O. umbonatus* and *F. favus*, the low ^{13}C group includes *Uvigerina* spp., *S. interrupta*, and *C. subglobosa*, the medium ^{13}C group includes *L. pauperata* and *N. umbonifera*, and the high ^{13}C group is made up of *P. wuellerstorfi* and *Q. auberiana*. Again, the aragonitic *H. elegans* is anomalous, with very high $\delta^{13}C$ values.

The carbon ranking's apparent independence of the oxygen ranking, may indicate that at least two mechanisms of fractionation are involved which operate more or less independently to produce the 'vital effects' in oxygen and carbon isotopic compositions. As was shown for oxygen, ranks apparently are time- and fertility-invariant (columns 1 compared with 2; and 1, 3, 4 compared with 5). Comparison of the degree of disequilibrium for carbon in the western area with that in the eastern area (Fig. 1*b*) shows that the change in disequilibrium is not as coherent as it seems for oxygen. However, most of the $\delta^{13}C$ values are closer to equilibrium in the eastern core than in the western cores—a reverse trend to the oxygen trend.

Interspecific and intraspecific ranges

In the Holocene samples, interspecific Δδ ranges are near 1‰ or less for oxygen and more than 1‰ for carbon (Table 2, Fig. 1). Intraspecific (within species) ranges in both $\delta^{18}O$ and $\delta^{13}C$ are commonly of the order of several tenths of 1‰ (also see ref. 10). Maximum intraspecific ranges within single samples are near 1.5‰ for both $\delta^{18}O$ and $\delta^{13}C$. Isolated determinations should, therefore, be interpreted with caution as indicators of changes in either ambient conditions or vital effects.

From Fig. 1, the relative position of *O. umbonatus* data with respect to the general trend of the other three species suggests that *O. umbonatus* may be less affected by fertility changes, although more data are necessary to confirm this. Nevertheless, our data do not support the conclusion of Belanger *et al.*[31] that *P. wuellerstorfi* is a more reliable indicator of bottom water than *O. umbonatus*.

Oxygen isotope record of the past 15,000 yr

The relationships between the isotopic signals of various deep-sea benthic foraminifera are further illustrated by a 15,000-yr record, based on closely spaced samples of Box Core ERDC-123 (Fig. 2). The carbon-14 stratigraphy of this core[32] (see Fig. 2, lower scale) shows that sedimentation was continuous. The sedimentation rate is ~3 cm kyr^{-1} from the base of the core to ~20 cm, and changes rather abruptly to 1.8 cm kyr^{-1} above 20 cm in the core, at ~10,000 yr BP. Sharply lower sedimentation rates for the post-Glacial, compared with the Glacial, have been reported for this area[23,33].

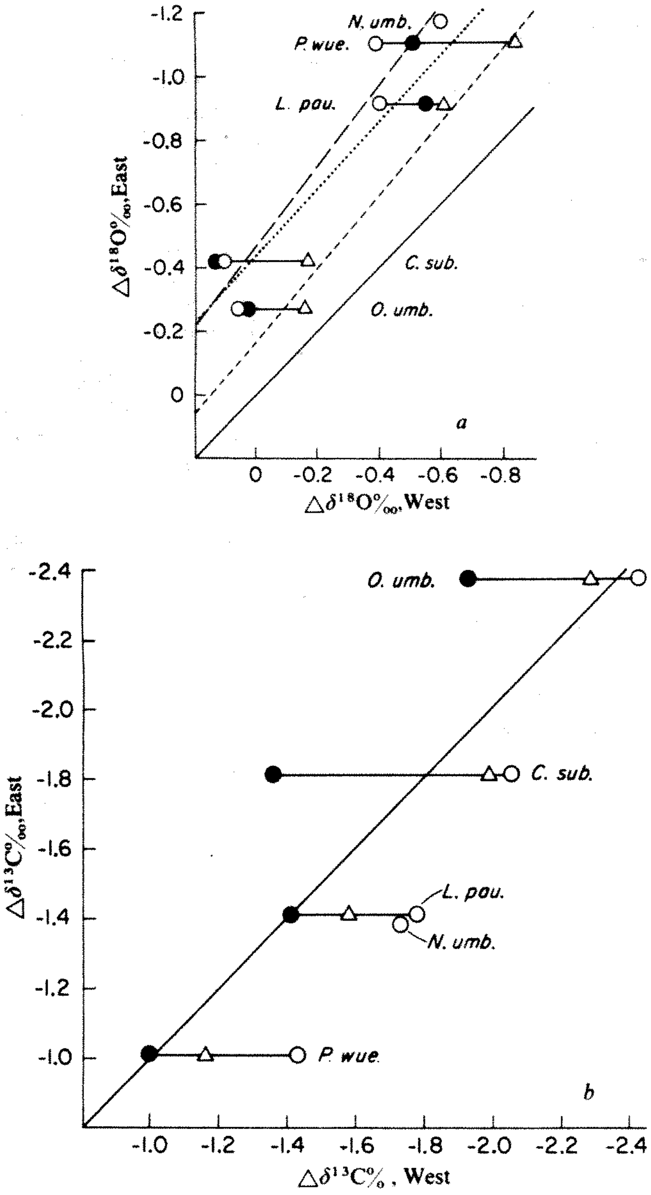


Fig. 1 Comparison of the degree of disequilibrium in recent benthic foraminifera for oxygen (*a*) and carbon (*b*) between west and east in the equatorial Pacific. $\Delta\delta$ values are from Table 2. Western cores: ERDC-92 (●), ERDC-123 (△), ERDC-135 (○); eastern core: PLDS-66. *N. umb.*, *Nuttalides umbonifera*; *P. wue.*, *Planulina wuellerstorfi*; *L. pau.*, *Laticarinina pauperata*; *C. sub.*, *Cassidulina subglobosa*; *O. umb.*, *Oridorsalis umbonatus*. The solid 45° line corresponds to an equal disequilibrium between west and east. The other lines are the best calculated fits for the oxygen isotopic data (long dashed line: ERDC-135, correlation coefficient of the fit, $r = 0.96$; dotted line: ERDC-92, $r = 0.93$; short dashed line: ERDC-123, $r = 0.99$).

The oxygen isotope record of shallow-living (*Globigerinoides sacculifer*) and deep-dwelling (*P. obliquiloculata* and *N. dutertrei*) planktonic foraminifera shows the familiar change from Glacial to post-Glacial values for this area[25]. The maximum rate of change occurs between 21 and 29 cm (13,000–10,000 ^{14}C yr). Three factors are involved in producing the change: (1) introduction of meltwater into the world ocean; (2) possible incomplete mixing of this meltwater; and (3) global warming. The relative importance of these factors to the rate of change is not clear, but we assume that the first can account for a total change of near 1‰, the second for a transient change of near 0.5‰, the third for about 0.4‰[24,34,35]. An additional factor, mixing by benthic organisms, decreases the recorded rate of change[36–39].

The oxygen isotope stratigraphy of benthic foraminifera also shows the deglaciation change, but continuing beyond 21 cm until values reach a light δ^{18}O plateau around 15 cm (7,500 ^{14}C yr BP). The difference in the isotope records does not necessarily indicate a 'lead-lag' relationship such as has been postulated for various Pleistocene palaeoceanographic signals[40,41]. The fact that the abundance of the carrier particles for the isotope signals changes across the Glacial–Holocene boundary will by itself change the shape of the signal, as pointed out earlier[34]. As the relative abundance of the benthic carrier decreases markedly across the boundary, its isotopic signal may lag behind that of the planktonic through the effects of mixing because proportionally more benthic signal carriers (with high δ^{18}O values) are mixed upwards when compared with the planktonic signal carriers. Recent numerical experiments by W. H. Hutson (personal communication) have demonstrated the importance of mixing on such lead-lag phenomena.

The various benthic signals are reasonably parallel, only the one of *F. favus* crosses any of the others. The rapid increase in δ^{18}O values in all four species near 10 cm (5,000 ^{14}C yr BP) and the distinct trend to lighter values near the surface of the sediment are remarkable. (The uppermost sediment in this core is in fact ^{14}C young and not yet absorbed into the mixed layer[32].)

The range of the deglacial δ^{18}O change, including the early Holocene plateau, is greatest in *P. wuellerstorfi*, the most oxygen-disequilibrated species of the four. Its range is 2‰. The other species show ranges between 0.9 and 1.4‰, for an average of 1.2‰. The cause of the benthic oxygen isotope plateau between 5,500 and 8,500 ^{14}C yr BP is not immediately clear, but it is possibly associated with the period of the "Climatic Optimum"[42].

We suggest the following interpretation of the oxygen isotope record: (1) a lag in the effect of glacial melting on the bottom water composition, with bottom waters reaching Holocene values ~2,000 yr after the upper waters; (2) a period in which benthic foraminifera have abnormally low ^{18}O values because of a decrease in bottom water production, and hence warming of bottom waters, during the climatic optimum; and (3) an increase of bottom water production after the climatic optimum, with a concomitant rapid drop in abyssal temperatures with the onset of "Neoglaciation"[42]. Future work will show whether these suggestions are acceptable or not. Clearly, however, ice volume cannot be read directly from the benthic signals (as has been variously proposed), because of: (1) differences in amplitudes between species; (2) existence of transient events throughout deglaciation and the Holocene; (3) mixing effects; and (4) intraspecific variation in δ^{18}O.

Carbon isotope record of the past 15,000 yr

The carbon isotope stratigraphy of the deep-dwelling planktonic foraminifera *P. obliquiloculata* is remarkable for its lack of structure. Apparently we need not postulate significant variation in δ^{13}C of upper ocean waters during the period in question.

The surface-dwelling *Globigerinoides sacculifer* shows a stepped signal: low values (1.65‰±0.15) from 15,000 to 9,000 ^{14}C yr ago, high values (2.0‰±0.2) since 9,000 yr BP. The change correlates with the previously mentioned changes in sedimentation rate and in benthic foraminiferal number. A distinct decrease in equatorial upwelling within the early Holocene, with concomitant stabilization of a ^{13}C-rich surface layer is suggested. A marked decrease in fertility from Glacial to post-Glacial time is also indicated by the δ^{13}C values of the fine sand fraction ('Fines'). Low δ^{13}C in this fraction is believed to indicate high metabolic activity[26], and such activity seems to decrease with time.

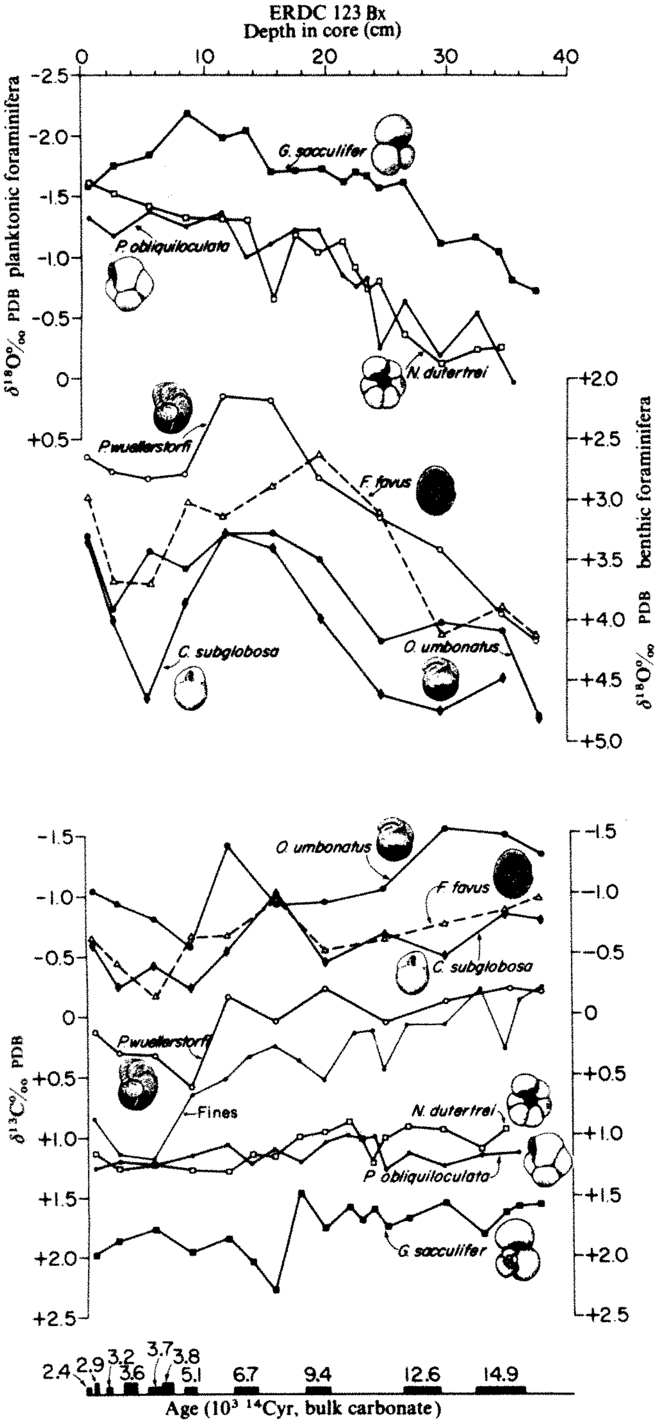


Fig. 2 Oxygen and carbon isotope stratigraphies in Box Core ERDC-123 for the shallow-living planktonic foraminifera *Globigerinoides sacculifer* (250–420 μm), the deep-dwelling ones *Pulleniatina obliquiloculata* and *Neogloboquadrina dutertrei* (250–420 μm) and the benthic species *Cassidulina subglobosa, Favocassidulina favus, Oridorsalis umbonatus* and *Planulina wuellestorfi* (>250 μm). Carbon isotope stratigraphy for the fine sand fraction ('Fines': 63–125 μm). Foram shells drawn by F. L. Parker (planktonics) and W. T. Coulbourn (benthics).

The decrease in fertility from Glacial to post-Glacial time is also reflected in an increase in δ^{13}C values of benthic foraminifera in post-Glacial sediments. Possibly the oxygen isotope plateau in the early Holocene is accompanied by a lightening of the ambient δ^{13}C values: if the plateau was partly due to warming, and hence incomplete mixing, isotopically light CO_2 should have built up. A CO_2 buildup would obscure a general δ^{13}C increase from lowered fertility, and shift the high benthic δ^{13}C values into the late Holocene. There is some indication in all four benthic signals that such a shift is indeed present, compared with the planktonic signal. Again, the effects of mixing on the signal, through changing signal carrier abundance may bias results in this direction.

Five of the benthic taxa analysed in this study (*C. subglobosa, F. favus, O. umbonatus, P. wuellerstorfi* and *Uvigerina* spp.) maintain the same relative positions in terms of δ^{18}O and δ^{13}C in Upper Miocene sediments[43], demonstrating that the relative intensity of vital effects in the various species tend to be retained over millions of years.

Conclusion

We have established the degree of disequilibrium precipitation in various deep-sea benthic foraminifera, and shown how this helps the interpretation of isotope record. Stratigraphies using multiple species are necessary for a more than superficial understanding of palaeoceanographic history, especially on a time scale of <10,000 yr. On this time scale, unfortunately, mixing processes disrupt the deep-sea record sufficiently to make stratiagraphic interpretations rather tentative, until more cores are studied and mixing processes are more quantitatively understood.

We thank Susan K. Culp, Richard F. Johnson, and Frances L. Parker for helpful discussions, and Thomas J. Walsh for assistance in collecting and sampling the box cores. We also thank Samuel M. Savin for constructive criticism. This study was supported by the NSF (Oceanography Section) grant OCE78-25587.

Received 2 July; accepted 2 December 1980.

1. Emiliani, C. *J. Geol.* **63,** 538–578 (1955).
2. Emiliani, C. *Science* **119,** 853–855 (1954).
3. Duplessy, J. C., Chenouard, L. & Villa, F. *Science* **188,** 1208–1209 (1975).
4. Streeter, S. S. & Shackleton, N. J. *Science* **203,** 168–171 (1979).
5. Shackleton, N. J. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 401–427 (Plenum, New York, 1977).
6. Shackleton, N. J. *Nature* **215,** 15–17 (1967).
7. Duplessy, J. C., Lalou, C. & Vinot, A. C. *Science* **168,** 250–251 (1970).
8. Vinot-Bertouille, A. C. & Duplessy, J. C. *Earth planet. Sci. Lett.* **18,** 247–257 (1973).
9. Vincent, E., Killingley, J. S. & Berger, W. H. *Geol. Soc. Am., Abstr. Prog.* **11,** 532–533 (1979).
10. Woodruff, F., Savin, S. M. & Douglas, R. G. *Mar. Micropaleontol.* **5,** 3–11 (1980).
11. Smith, P. B. & Emiliani, C. *Science* **160,** 1335–1336 (1968).
12. Shackleton, N. J. *Colloques Int. Cent. natn. Rech. scient.* **219,** 203–209 (1974).
13. Boersma, A. & Shackleton, N. J. *Init. Rep. DSDP* Leg **39,** 911–924 (1974).
14. Duplessy, J. C. in *Climatic Changes* (ed. Gribbin, J.) 46–67 (Cambridge University Press, 1978).
15. Pastouret, L., Chamley, H., Delibrias, G. & Duplessy, J. C. *Oceanol. Acta* **1,** 217–232 (1978).
16. Keigwin, L. D. Jr *Earth planet. Sci. Lett.* **45,** 361–382 (1979).
17. Keigwin, L. D. Jr, Bender, M. L. & Kennett, J. P. *Science* **205,** 1386–1388 (1979).
18. Shackleton, N. J. & Cita, M. B. *Init. Rep. DSDP* Leg **47,** 433–445 (1979).
19. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3,** 39–55 (1973).
20. Berger, W. H., Johnson, T. C. & Hamilton, E. L. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 543–567 (Plenum, New York, 1977).
21. Johnson, T. C., Hamilton, E. L. & Berger, W. H. *Mar. Geol.* **24,** 259–277 (1977).
22. Parker, F. L. & Berger, W. H. *Deep Sea Res.* **18,** 73–107 (1971).
23. Peng, T. H., Broecker, W. S. & Berger, W. H. *Quat. Res.* **11,** 141–149 (1979).
24. Berger, W. H. & Killingley, J. S. *Science* **197,** 563–566 (1977).
25. Culp, S. K. M.S. thesis, Univ. Hawaii (1977).
26. Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanol. Acta* **1,** 203–216 (1978).
27. Bandy, O. L. *J. Sediment. Petrol.* **24,** 60–61 (1954).
28. Blackmon, P. D. & Todd, R. *J. Paleontol.* **33,** 1–15 (1959).
29. Vénec-Peyré, M. T. & Jaeschke-Boyer, H. *C.r. hebd. Séanc. Acad. Sci, Paris* **287,** 607–609 (1978).
30. Williams, D. F., Sommer, M. A. & Bender, M. L. *Earth planet. Sci. Lett.* **36,** 391–403 (1977).
31. Belanger, P. E., Curry, W. B. & Matthews, R. K. in *Oxygen and Carbon Isotopes in Foraminifera* (eds Berger, W. H., Bé, A. W. H. & Vincent, E.) (Elsevier, Amsterdam, in the press).
32. Krishnamurthy, R. V., Lal, D. Somayajulu, B. L. K. & Berger, W. H. *Proc. Ind. Acad. Sci.* 88A Pt. II, 3, 273–283 (1979).

33. Berger, W. H. & Johnson, R. F. *Earth planet. Sci. Lett.* **41,** 223–227 (1978).
34. Berger, W. H., Johnson, R. F. & Killingley, J. S. *Nature* **269,** 661–663 (1977).
35. Vincent, E. & Berger, W. H. in *The Sea*, 7 (ed., Emiliani, C.) (Wiley, New York, in the press).
36. Berger, W. H. & Heath, G. R. *J. mar. Res.* **26,** 134–143 (1968).
37. Nozaki, Y., Cochran, J. K., Turekian, K. K. & Keller, G. *Earth planet. Sci. Lett.* **34,** 167–184 (1977).
38. Peng, T. H., Broecker, W. S., Kipphut, G. & Shackleton, N. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 355–373 (Plenum, New York, 1977).
39. Berger, W. H., Ekdale, A. A. & Bryant, P. F. *Mar. Geol.* **32,** 205–230 (1979).
40. Moore, T. C. Jr, Pisias, N. G. & Heath, G. R. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 145–165 (Plenum, New York, 1977).
41. Thompson, P. R. *J. Foram. Res.* **6,** 208–227 (1976).
42. Flint, R. F. *Glacial and Quaternary Geology* (Wiley, New York, 1971).
43. Vincent, E., Killingley, J. S. & Berger, W. H. *Mar. Micropaleonol.* **5,** 185–203 (1980).
44. Epstein, S., Buchsbaum, K., Lowenstam, H. & Urey, H. C. *Bull. geol. Soc. Am.* **64,** 1315–1326 (1952).
45. Reid, J. L. *Intermediate Waters of the Pacific Ocean*, 7–85 (The Johns Hopkins University Press, 1965).
46. Craig, H. & Gordon, L. I. in *2nd Conf., Stable Isotopes in Oceanographic Studies and Paleotemperatures* (ed. Tongiorgi, E.) 9–130 (Consiglio Nazionale delle Richerche, Rome, 1965).
47. Emrich, K., Ehhalt, D. H. & Vogel, J. C. *Earth planet. Sci. Lett.* **8,** 363–371 (1970).
48. Kroopnick, P. M., thesis, Univ. California, San Diego (1971).

Mouse liver and salivary gland α-amylase mRNAs differ only in 5′ non-translated sequences

Otto Hagenbüchle, Mario Tosi, Ueli Schibler, Raymonde Bovey, Peter K. Wellauer & Richard A. Young*

Swiss Institute for Experimental Cancer Research, Chemin des Boveresses, 1066 Epalinges, Switzerland

The sequence of the 1,773-nucleotide major and 1,806-nucleotide minor mouse liver α-amylase mRNAs differ only with respect to approximately 30 additional residues at the extreme 5′ terminus of the minor species. Comparison of the liver α-amylase mRNAs with their salivary gland counterpart reveals that these mRNAs share identical coding and 3′-noncoding sequences, but contain distinct 5′-terminal residues. These data suggest that all three mRNAs might be transcribed from the same gene.

TO gain insight into the mechanisms which govern tissue-specific gene expression, we have investigated α-amylase mRNA biosynthesis in various mouse tissues. Early reports from other investigators indicated that α-amylase is produced in the pancreas, the salivary gland and the liver[1-3] and that the pancreatic protein differs from the other two with respect to antigenic and charge properties[4-9]. In our recent report of the primary structure of the α-amylase mRNAs which accumulate in the pancreas and the salivary gland, we found that the two mRNAs differ in nucleotide sequence, indicating that they are transcribed from different genes[10,11].

The liver α-amylase mRNA is less well characterized. Schibler *et al.*[10] showed that this mRNA is 100-fold less abundant and ~100 nucleotides longer than its salivary gland counterpart. However, the two species were found to have very similar sequences as indicated by their melting behaviour and resistance to S_1 nuclease digestion when part of α-amylase cDNA–mRNA hybrids. Restriction endonuclease digestion of cloned cDNAs made against liver and salivary gland α-amylase mRNA revealed no differences in cleavage sites for the various enzymes tested[12].

To elucidate structural relationships between α-amylase mRNAs produced in the salivary gland and liver in more detail, we have determined the nucleotide sequence of liver α-amylase mRNA. Here we report the intriguing finding that the liver and salivary gland α-amylase mRNAs encode identical proteins but differ in the terminal portion of their 5′-non-translated regions.

5′-terminal sequences account for the length difference between salivary gland and liver α-amylase mRNAs

Hybrids between salivary gland α-amylase cDNA and liver mRNA are resistant to S_1 nuclease along the entire length of the 1,600-nucleotide cDNA[10], suggesting that sequences which are responsible for the length difference between salivary gland and liver mRNAs reside at one or both ends of the liver species. To locate these extra sequences, we exploited the ability of RNase H to digest the RNA moiety of RNA–DNA hybrids[13-15]. The experimental approach is outlined in Fig. 1*a*. The 520-base pair *Eco*RI fragment of the cDNA clone pMSa104 (Fig. 1*b*), which contains translated sequences from the salivary gland α-amylase mRNA, was hybridized to polyadenylated RNA from pancreas, salivary gland or liver. Hybrids were then digested with RNase H and the products separated on a methyl mercury hydroxide agarose gel and transferred to diazotized (DMB) paper. The 5′ or 3′ RNase H-resistant RNA fragments were visualized by hybridization with appropriate 5′- or 3′-specific probes generated from the pMSa104 cDNA clone (Fig. 1*b*).

The autoradiographs shown in Fig. 1 indicate that sequences which account for the differences in size between the liver and salivary gland α-amylase mRNAs reside at the 5′ terminus. The results for the pancreas and salivary gland mRNAs (Fig. 1*c*) concur with those expected: only one major α-amylase mRNA accumulates in each of these two tissues and the salivary gland mRNA contains 80 more 5′-terminal noncoding residues than its pancreatic counterpart[10,11]. Two 5′-terminal liver RNA fragments are revealed in this experiment. Using the predicted sizes of the pancreas and salivary gland 5′-terminal fragments, we estimate that the two liver fragments are 490 and 520 nucleotides long. Densitometer tracings indicate that the larger fragment is present at 20% the level of the smaller one. For convenience, we will refer to these two liver RNAs as 5′-major and 5′-minor RNAs, reflecting their relative concentrations.

No apparent length differences are observed in 3′-terminal α-amylase mRNA fragments from the three tissues (Fig. 1*d*). However, an overexposure of the autoradiograph revealed a minor RNA fragment in both salivary gland and liver; Fig. 1*e* shows a longer exposure of the lane containing liver RNA. The size and abundance of this fragment indicate that it originates from a previously described minor α-amylase mRNA species (3′-minor RNA). Present at about 5% the level of total α-amylase mRNA in both tissues, the 3′-minor species contains major species sequences plus 237 additional nucleotides at its 3′ end[12].

---

[*] Present address: Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305.

0028–0836/81/080643—04$01.00

Sequence analysis of liver α-amylase mRNAs

To determine the liver α-amylase mRNA sequence, our strategy included elucidating the cDNA sequence. Double-stranded cDNA was synthesized against liver poly(A)$^+$ RNA and cloned in *E. coli*[12]. Due to the low abundance of α-amylase mRNA in this tissue (0.02% of poly(A)$^+$ RNA), only one recombinant plasmid (pMLa52) was obtained. Restriction endonuclease mapping of pMLa52 DNA suggested that it contained sequences complementary to the 3′-minor α-amylase mRNA; subsequent sequence analysis of residues representing the extreme 3′ terminus confirmed this interpretation[12]. Because the 3′-minor species is simply an extended version of the 3′-major liver α-amylase mRNA (see above), we could use pMLa52 DNA to determine sequences for both mRNAs. Therefore, the entire sequence of the cDNA of pMLa52 was determined according to Maxam and Gilbert[16] as outlined in Fig. 2 legend. Nucleotides 518–1,773 of the liver α-amylase mRNA (Fig. 3) were elucidated from pMLa52. These residues are identical to those determined for the salivary gland α-amylase mRNA[11]. As 3′-minor nucleotide sequences have been reported elsewhere[12], only those for the 3′-major species are included in Fig. 3.

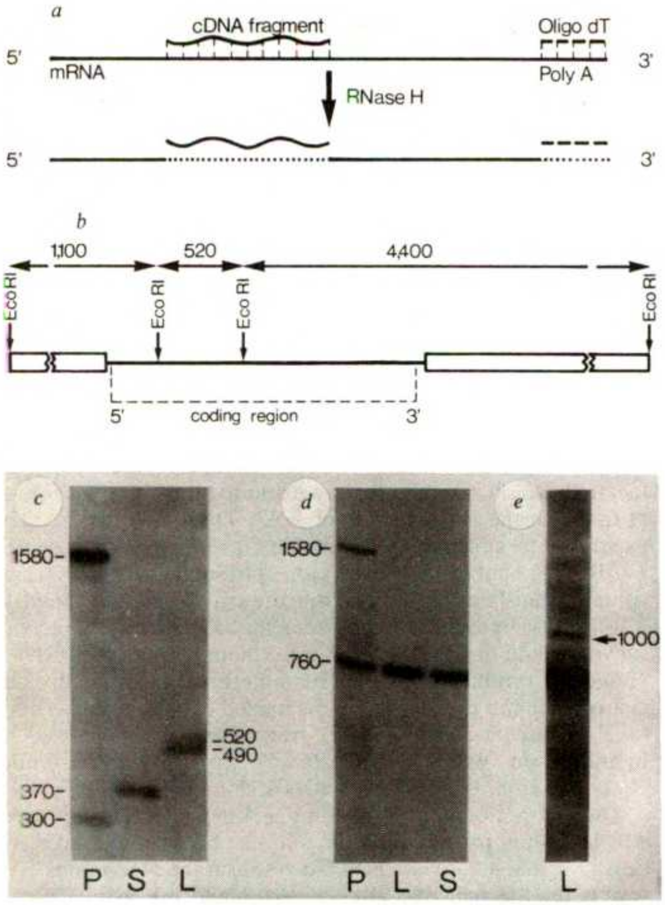


Fig. 1 Locating sequences which account for the length differences between liver, salivary gland and pancreas α-amylase mRNAs. *a*, the strategy used to generate 5′- and 3′-terminal fragments of α-amylase mRNA using RNase H. *b*, Diagram of pMSa104 DNA[10] linearized with respect to the pBR322 *Eco*RI site. Salivary gland α-amylase cDNA insert is represented as a line, pBR322 vector is boxed. Arrows delineate the *Eco*RI fragments used in this experiment. Numbers are in base pairs. *c*, *d*, Autoradiographs of DBM paper-bound RNase H-resistant RNA fragments probed with 5′-end-specific (1.1 kilobase *Eco*RI fragment from pMSa104; *c*) or 3′-end-specific (4.4 kilobase *Eco*RI fragment from pMSa104; *d*) labelled DNA restriction fragments (see *b*). Poly(A)$^+$ RNA from pancreas (P, 500 ng), salivary gland (S, 1 μg) and liver (L, 40 μg) was hybridized to 100 ng (300 ng for pancreas) of the 520-base pair *Eco*RI restriction fragment from pMSa104 (*b*) in a final volume of 100 μl of 80% formamide, 0.05% SDS, 0.4 M NaCl, 1 mM EDTA, 10 mM PIPES, *p*H 6.4, for 3 h at 45 °C. After ethanol precipitation, the mRNA/DNA hybrids were digested with RNase H for 15 min at 37 °C in 10 mM Tris-HCl *p*H 7.6, 0.13 M NH_4Cl, 0.01 M Mg acetate, 5% (w/v) sucrose, 20 $\mu g\ ml^{-1}$ oligo(dT), and 1 unit RNase H (Enzo) in a final volume of 30 μl (8 units RNase H in 300 μl for liver). After phenol extracting and ethanol precipitating, equivalent amounts of RNase H-resistant α-amylase mRNA sequences were electrophoresed on a 2% methyl mercury hydroxide agarose gel, transferred to DBM paper[10,20] and hybridized with the 5′- or 3′-end-specific probes. Numbers are in nucleotides, the 1,580-nucleotide bands in lane P represent uncleaved, deadenylated α-amylase mRNA. *e*, A 20-fold longer exposure of lane L in *d*. An arrow indicates the minor species fragment.

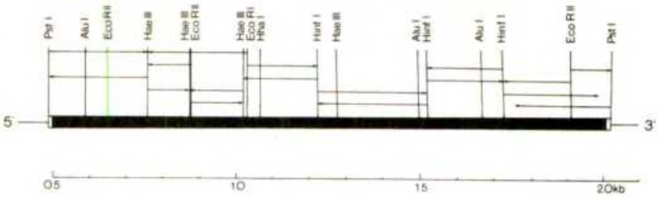


Fig. 2 Restriction map and strategy used to determine the nucleotide sequence of the liver α-amylase cDNA insert of recombinant plasmid pMLa52. Restriction endonuclease cleavage sites were determined by the technique of Smith and Birnstiel[21]. 5′-end-labelled fragments were sequenced according to Maxam and Gilbert[16]. The direction and extent of sequence information determined for each fragment examined are indicated by an arrow. The black box represents mRNA sequences, the open boxes indicate poly(dG-dC) generated during the cloning procedure. kb, kilobases.

To obtain primary structure information for 5′-terminal sequences not included in the cDNA clone, we used reverse transcriptase-catalysed primer extension as described by Hagenbüchle *et al.*[11]. A 50-nucleotide single-stranded DNA fragment, isolated from the salivary gland α-amylase cDNA clone pMSa104 and containing nucleotide residues 191–241 (Fig. 3), was 5′-end-labelled according to Maxam and Gilbert[16] and hybridized to liver polyadenylated RNA. Two cDNA transcripts were produced after incubation with AMV reverse transcriptase which differed by approximately 30 nucleotides. The sizes and relative amounts of the two cDNAs were consistent with those predicted by the RNase H sizing experiment for the two liver mRNAs (Fig. 1). A unique sequence was elucidated for the 5′-major species cDNA (residues 1–191, Fig. 3) and simultaneous determination of the cytosine and purine residues of the 5′-minor species cDNA revealed that it is an extended version of the 5′-major species (data not shown). Notably, nucleotides 159–191 (Fig. 3) are identical to those determined for the comparable portion of the salivary gland α-amylase mRNA. Due to the low cellular concentration of liver α-amylase mRNA, direct RNA sequencing by chemical decapping and 5′-end labelling was not feasible; therefore, the 5′-terminal few residues could not be determined.

Sequences determined for the liver α-amylase cDNA produced by primer extension (nucleotides 1–191, Fig. 3) do not overlap those elucidated from pMLa52 DNA (nucleotides 518–1,773, Fig. 3). Of these, residues 159–191 and 518–1,773 are identical to those determined for the salivary gland α-amylase mRNA. We assume that the two α-amylase mRNAs also share residues 192–517 (sequences not examined for the liver species) because S_1 nuclease digestion and thermal melting curves of liver α-amylase mRNA/salivary gland α-amylase cDNA hybrids reveal no sequence heterogeneity[10].

Features of the liver α-amylase mRNAs

The results obtained by a combination of RNase H digestion, primed cDNA synthesis and sequence analysis indicate that several α-amylase mRNA species accumulate in the liver (Fig. 4). The 1,773-nucleotide major mRNA (5′-major/3′-major) accounts for 80% of the total α-amylase mRNA population. The remaining 20% are composed of 5′-minor species which contain ~30 additional nucleotides at their extreme 5′ termini. Among these two mRNA species, 5% contain the 237 3′-minor species residues; the distribution of 3′-minor species among 5′-major and 5′-minor mRNAs has not been determined. We do not know whether these minor species serve some special function or whether they represent products of alternative synthesis or processing pathways.

```
L (m7Gppp)XCCAUCAUAGGUCACUGUGGAGCUCAGAUCACAGUGCUGACAGAAUCCAUAUUUGGAGAAUUACAUAAGGUUUGAAAGAGAGGAUAGUGAAAGGAUACG   100

   AAUUCCUAAAAAACGUUUAAUCUGGCCUUUUGUUUGAACGAAAGAGAAAUUGAAACCAA
                                                              AUGAAAUAAAUUAGUUGUUAGAAAGAAUACUGCCAACAGCAU   200
   S (m7Gppp)AUUCUCAACGUAAAUCAGAAGAUUCCAACUGCAGUGGAAGGCAGCAC

   AGCAAAUGAAAUUCUUCCUGCUGCUUUCCCUCAUUGGAUUCUGCUGGGCCCAAUAUGACCCACAUACUCAAUAUGGACGAACUGCUAUUAUCCACCUGU   300
         MetLysPhePheLeuLeuLeuSerLeuIluGlyPheCysTrpAlaGlnTyrAspProHisThrGlnTyrGlyArgThrAlaIluIluHisLeuP

   UUGAGUGGCGCUGGGUUGAUAUUGCUAAGGAAUGUGAGAGAUACUUAGCUCCUAAUGGAUUUGCAGGUGUGCAGGUCUCUCCACCCAAUGAAAACAUCGU   400
   heGluTrpArgTrpValAspIluAlaLysGluCysGluArgTyrLeuAlaProAsnGlyPheAlaGlyValGlnValSerProProAsnGluAsnIluVa

   AGUCCACAGCCCUUCAAGACCAUGGUGGGAAAGAUAUCAACCAAUUAGCUACAAAAUAUGUUCCAGGUCUGGAAAUGAAGAUGAAUUCAGGGACAUGGUG   500
   lValHisSerProSerArgProTrpTrpGluArgTyrGlnProIluSerTyrLysIluCysSerArgSerGlyAsnGluAspGluPheArgAspMetVal

   AACAGGUGCAACAAUGUUGGUGUCCGUAUUUAUGUGGAUGCUGUCAUUAACCACAUGUGUGGAGUGGGGGCUCAAGCUGGACAAAGCAGUACAUGUGGAA   600
   AsnArgCysAsnAsnValGlyValArgIluTyrValAspAlaValIluAsnHisMetCysGlyValGlyAlaGlnAlaGlyGlnSerSerThrCysGlyS

   GUUAUUUCAACCCAAAUAACAGGGACUUUCCUGGAGUUCCCUAUUCUGGUUUUGACUUUAAUGAUGGAAAAUGUAGAACUGCAAGUGGAGGUAUCGAGAA   700
   erTyrPheAsnProAsnAsnArgAspPheProGlyValProTyrSerGlyPheAspPheAsnAspGlyLysCysArgThrAlaSerGlyGlyIluGluAs

   CUACCAAGAUGCUGCUCAGGUCAGAGAUUGUCGUCUGUCUGGCCUUCUGGAUCUUGCACUUGAGAAAGAUUAUGUUCGAACCAAGGUGGCUGACUAUAUG   800
   nTyrGlnAspAlaAlaGlnValArgAspCysArgLeuSerGlyLeuLeuAspLeuAlaLeuGluLysAspTyrValArgThrLysValAlaAspTyrMet

   AACCAUCUCAUUGACAUUGGCGUAGCAGGGUUCAGACUUGAUGCUUCUAAGCACAUGUGGCCUGGAGACAUAAAGGCAAUUUUGGACAAACUGCAUAAUC   900
   AsnHisLeuIluAspIluGlyValAlaGlyPheArgLeuAspAlaSerLysHisMetTrpProGlyAspIluLysAlaIluLeuAspLysLeuHisAsnL

   UCAAUACAAAAUGGUUCUCCCAAGGAAGCAGACCUUUCAUUUUCCAAGAGGUGAUUGAUCUGGGUGGUGAGGCAGUGUCAAGUAAUGAGUAUUUUGGAAA   1,000
   e  nThrLysTrpPheSerGlnGlySerArgProPheIluPheGlnGluValIluAspLeuGlyGlyGluAlaValSerSerAsnGluTyrPheGlyAs

   UGGCCGUGUGACAGAAUUCAAAUAUGGAGCAAAAUUGGGCAAAGUUAUGCGCAAGUGGGAUGGAGAAAAGAUGUCCUACUUAAAGAACUGGGGAGAAGGU   1,100
   nGlyArgValThrGluPheLysTyrGlyAlaLysLeuGlyLysValMetArgLysTrpAspGlyGluLysMetSerTyrLeuLysAsnTrpGlyGluGly

   UGGGGUUUGAUGCCUUCUGACAGAGCCCUUGUGUUUGUGGACAACCAUGACAAUCAGCGAGGACAUGGUGCUGGGGGAGCAUCCAUCUUGACAUUCUGGG   1,200
   TrpGlyLeuMetProSerAspArgAlaLeuValPheValAspAsnHisAspAsnGlnArgGlyHisGlyAlaGlyGlyAlaSerIluLeuThrPheTrpA

   AUGCUAGACUCUAUAAAAUGGCUGUUGGCUUUAUGUUGGCUCAUCCUUAUGGUUUCACACGGGUGAUGUCAAGUUACUAUUGGCCAAGAAAUUUCCAGAA   1,300
   spAlaArgLeuTyrLysMetAlaValGlyPheMetLeuAlaHisProTyrGlyPheThrArgValMetSerSerTyrTyrTrpProArgAsnPheGlnAs

   UGGAAAAGAUGUCAAUGACUGGGUUGGACCACCAAAUAACAAUGGAAAAACCAAAGAAGUGAGCAUUAACCCAGACAGCACUUGUGGCAAUGACUGGAUC   1,400
   nGlyLysAspValAsnAspTrpValGlyProProAsnAsnAsnGlyLysThrLysGluValSerIluAsnProAspSerThrCysGlyAsnAspTrpIlu

   UGUGAACAUCGAUGGCGUCAAAUAAGGAACAUGGUUGCCUUCAGAAAUGUCGUCAAUGGUCAGCCUUUUGCAAACUGGUGGGAUAAUGACAGCAACCAGG   1,500
   CysGluHisArgTrpArgGlnIluArgAsnMetValAlaPheArgAsnValValAsnGlyGlnProPheAlaAsnTrpTrpAspAsnAspSerAsnGlnV

   UAGCUUUUGGCAGAGGAAACAAAGGACUCAUUGUCUUUAACAAUGAUGACUGGGCUUUGUCAGAAACUUUACAGACUGGUCUUCCUGCUGGCACAUACUG   1,600
   alAlaPheGlyArgGlyAsnLysGlyLeuIluValPheAsnAsnAspAspTrpAlaLeuSerGluThrLeuGlnThrGlyLeuProAlaGlyThrTyrCy

   UGAUGUCAUUUCUGGAGAUAAAGUCGAUGGCAAUUGCACUGGAAUAAAAGUCUAUGUUGGCAAUGAUGGCAAAGCUCACUUUUCUAUUAGUAACUCUGCC   1,700
   sAspValIluSerGlyAspLysValAspGlyAsnCysThrGlyIluLysValTyrValGlyAsnAspGlyLysAlaHisPheSerIluSerAsnSerAla

   GAAGACCCAUUUAUUGCAAUCCAUGCAGAGUCAAAAAUAUAAAUUUAAAAUAAACACAUAUUGAGAGCAUCAn
   GluAspProPheIluAlaIluHisAlaGluSerLysIlu
```

Fig. 3 Comparison of the nucleotide sequence of the major liver α-amylase mRNA and its salivary gland counterpart. Numbering begins with nucleotide X (which represents the few initial nucleotides that were not determined, see text) in the liver mRNA. The nucleotide residues 112–158 shown for the salivary gland mRNA were taken from O.H. *et al.*[11] but contain two nucleotides (residues 113 and 114) that were excluded from the sequence reported earlier. The two mRNAs are assumed to be capped. S and L indicate the tissue-specific salivary gland and 5′-major liver α-amylase mRNA leaders. The first AUG codon is underlined; that actually used to initiate α-amylase synthesis is boxed, as is the terminator triplet. Both mRNAs predict the amino acid sequence shown below the nucleotide sequence.

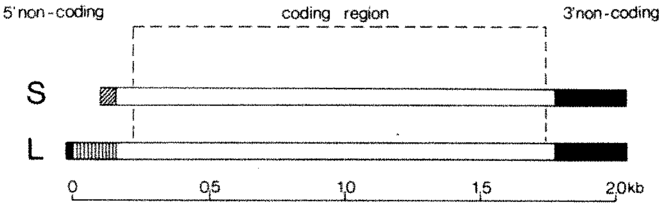


Fig. 4 Schematic comparison of the liver and salivary gland α-amylase mRNAs. The diagram is based on data presented here and elsewhere[11,12], as described in the text. The open boxes represent common sequences; the black boxes designate residues found in the minor salivary (S) and liver (L) α-amylase mRNA species. Tissue-specific leader residues are shaded diagonally for salivary gland mRNA or vertically for the liver mRNA species. Numbering is that in Fig. 3.

Tissue-specific 5'-noncoding segments

Figure 3 compares the salivary gland α-amylase mRNA sequence with that of its major liver counterpart. The two mRNAs share nucleotides 159–1,773, which include the entire coding and 3'-noncoding regions. These salivary gland mRNA sequences have been discussed in detail previously[11]. Briefly, translation of the α-amylase precursor can occur from only one coding frame (the two others contain many termination triplets); this frame predicts a precursor protein whose characteristics mirror those determined for α-amylase produced *in vivo* and *in vitro*[10,17].

In addition to the coding and 3'-noncoding regions, a portion of the 5'-non-translated region is identical in the liver and salivary gland α-amylase mRNAs. The 204- and 93-nucleotide 5'-terminal noncoding regions share the 48 residues immediately adjacent the coding region. This common segment (nucleotides 159–206, Fig. 3) has the intriguing feature that it begins with an AUG codon which is not used to initiate translation of α-amylase even though it is the one nearest the 5' terminus. The first AUG is followed by a lysine codon, then the termination triplet UAA. The AUG used to initiate translation[11] lies 14 codons downstream. As all other eukaryotic cellular mRNAs for which complete sequence information is available initiate translation at the AUG codon located nearest the 5' end[18], the salivary gland and liver α-amylase mRNAs are apparent exceptions. Whether the first AUG in these α-amylase mRNAs has some regulatory role in translation is unclear.

It is the initial 158 nucleotides of the major liver species and the 47 5'-terminal residues of its salivary gland counterpart that distinguish the α-amylase mRNAs which accumulate in these two tissues. Hence, these two species have distinct 5'-terminal noncoding regions but direct the translation of an identical α-amylase polypeptide. The presence of distinguishing 5'-non-translated sequences might be the consequence of tissue-specific requirements for the rate of translation initiation. Although the influence of particular sequences 5' to the AUG initiation codon is unclear[18], translation initiation efficiency *in vitro* can be affected by the length of the 5'-non-translated region (C. Queen, personal communication). Therefore, the different lengths of the 5'-non-translated segments in salivary gland and liver α-amylase mRNAs (93 and 204 residues, respectively) might influence their relative translatability. Finally, the distinct 5'-non-translated segments might also specify different lifetimes for these tissue-specific mRNAs.

The observation that liver and salivary gland α-amylase mRNAs share identical coding and 3'-noncoding residues suggests that they are transcribed from the same DNA sequence. Indeed, recent analysis of genomic DNA demonstrates that both liver and salivary gland α-amylase mRNAs are specified by a single DNA sequence in the mouse haploid genome[19]. Two different segments of DNA specify the distinct 5'-non-translated regions of these mRNAs. These findings have implications for the mechanisms which govern the expression of this α-amylase gene; for example, tissue-specific splice events must be responsible for the correct biosynthesis of α-amylase mRNA in the liver and salivary gland.

We thank B. Hirt for his support, S. Cherpillod for typing the manuscript, D. Petrelli for technical assistance, L. J. Korn and B. Meyer for comments on the manuscript, and J. Beard for supplying reverse transcriptase through the Office of Program Resources and Logistics, Viral Cancer Program (Viral Oncology Division of Cancer Cause and Prevention), NCI. This work was supported by grant 3.148.77 from the Swiss NSF.

Received 29 October 1980; accepted 13 January 1981.

1. Loyter, A. & Schramm, M. *Biochim. biophys. Acta* **65,** 200–206 (1962).
2. Arnold, M. & Rutter, W. J. *J. biol. Chem.* **238,** 2760–2765 (1963).
3. Hjorth, J. P., Meisler, M. H. & Nielsen, J. T. *Genetics* **92,** 915–930 (1979).
4. Messer, M. & Dean, R. I. *Biochem. J.* **151,** 17–22 (1975).
5. Sick, K. & Nielsen, J. T. *Hereditas* **51,** 291–296 (1964).
6. Hammerton, K. & Messer, M. *Biochim. biophys. Acta* **244,** 441–451 (1971).
7. Sanders, T. G. & Rutter, W. J. *Biochemistry* **11,** 130–136 (1972).
8. Kaplan, R. D., Chapman, V. & Ruddle, F. H. *J. Hered.* **64,** 155–157 (1973).
9. Takeuchi, T., Matsushima, T. & Sugimura, T. *Biochim. biophys. Acta* **403,** 122–130 (1975).
10. Schibler, U., Tosi, M., Pittet, A.-C. & Wellauer, P. K. *J. molec. Biol.* **142,** 93–116 (1980).
11. Hagenbüchle, O., Bovey, R. & Young, R. A. *Cell* **21,** 179–187 (1980).
12. Tosi, M., Young, R. A., Hagenbüchle, O. & Schibler, U. *Nucleic Acids Res.* (submitted).
13. Donis-Keller, H. *Nucleic Acids Res.* **7,** 179–192 (1979).
14. Robertson, H. D. & Dunn, J. J. *J. biol. Chem.* **250,** 3030–3056 (1975).
15. Sippel, A. E., Stavrianopoulos, J. G., Schütz, G. & Feigelson, P. *Proc. natn. Acad. Sci. U.S.A.* **71,** 4635–4639 (1974).
16. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).
17. Danielsson, A., Marklund, S. & Stigbrand, T. *Int. J. Biochem.* **6,** 585–589 (1975).
18. Kozak, M. *Cell* **15,** 1109–1123 (1978).
19. Young, R. A., Hagenbüchle, O. & Schibler, U. Cell (in the press).
20. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
21. Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3,** 2387–2398 (1976).

Control of vascular permeability by polymorphonuclear leukocytes in inflammation

Caroline V. Wedmore & T. J. Williams*

Department of Pharmacology, Institute of Basic Medical Sciences, Royal College of Surgeons of England, Lincoln's Inn Fields,
London WC2A 3PN, UK

Polymorphonuclear leukocytes infiltrate tissues in response to an inflammatory stimulus to remove invading micro-organisms and cell debris. We present evidence that these scavenging cells have another, more sophisticated role in that they are involved in the control of fluid efflux through the blood vessel wall which leads to tissue oedema.

SINCE the first observations of leukocytes accumulating at a site of inflammation, their relationship to changes occurring in blood vessels has been a source of controversy. Studies of animal models of inflammation (for example, rat paw oedema in response to carrageenin[1]) have suggested that acute vascular responses, vasodilatation and increased vascular permeability, result from the sequential release of low molecular weight mediators—histamine, serotonin, bradykinin and prostaglandins[2]. During these acute vascular changes polymorphonuclear (PMN) leukocytes, predominantly neutrophils, accumulate slowly in the tissues reaching significant numbers after several hours. Based on these models, and allergic models such as the Arthus reaction[3], several mechanisms have been proposed whereby PMN leukocytes could contribute to oedema. The most widely accepted of these, first postulated by Metchnikoff[4], was that during phagocytosis PMN leukocytes release substances (lysosomal enzymes[5] for example) which damage vascular endothelial cells—the cells lining the blood vessel wall which are normally responsible for retaining plasma proteins within the lumen. We suggest an alternative hypothesis, that PMN leukocytes, responding to a chemical signal in the tissue, can control the permeability of the blood vessel wall to plasma proteins. Three such signalling substances are described here. Although PMN leukocytes are present in high numbers in tissues several hours after the initiation of an inflammatory response, the evidence suggests that these cells interact with vascular endothelial cells to regulate permeability within a few minutes of the inflammatory stimulus.

---

* To whom correspondence should be addressed.

The chemical mediators of inflammatory oedema in man are unknown; there is little evidence that histamine, serotonin and bradykinin are important. We suggest that the mechanisms involving leukocytes described here may be more relevant.

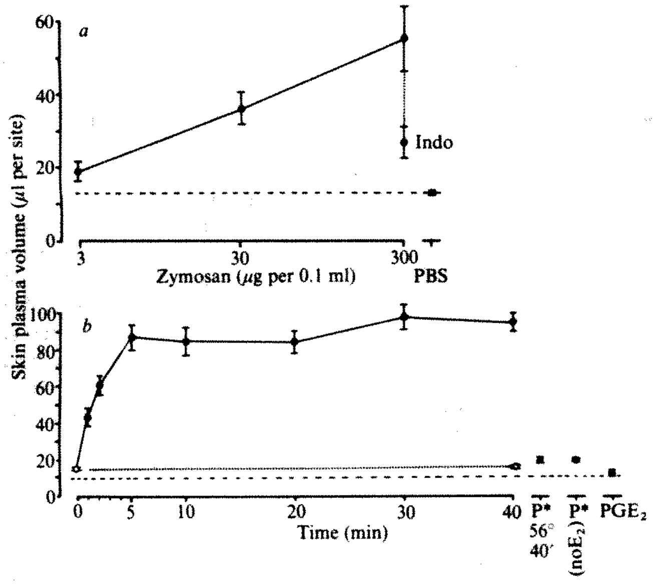


Fig. 1 *a*, Dose–response curve of oedema induced in rabbit skin by intradermal injection of zymosan. *b*, Time course of generation of permeability-increasing activity in blood plasma incubated with zymosan at 37°C. In all experiments plasma exudation was measured using the accumulation of intravenously injected ^{131}I-human serum albumin in rabbit (male NZW, 3–3.5 kg) dorsal skin in response to 0.1-ml intradermal injections of the agents. After a fixed period (30 min in all experiments except where stated) animals were killed, the skin was removed, the injection sites were excised (16-mm diameter punch) and counted in a γ-counter. *a* Shows the increase in exudate volume (measured over a 90-min period) in response to increasing amounts of injected zymosan. Indomethacin (Indo), an inhibitor of prostaglandin synthesis, injected locally at 0 min and 60 min (3×10^{-9} mol per site) suppressed exudation, as shown, an effect which was reversed by exogenous prostaglandin[14]. Oedema induced by intradermally injected zymosan was also suppressed by nitrogen mustard (1.75 mg per kg intravenous injection, 4 days before testing in the skin) which depletes circulating PMN leukocytes. Plasma volumes (90 min accumulation) of skin sites injected with 100 μg of zymosan were 40.0±6.3 μl in control animals, and 13.1±2.1 μl in depleted animals ($n=4$ rabbits, saline control (PBS, phosphate-buffered saline) values 16.6±1.8 μl and 15.7±2.0 μl, respectively). Unlike suppression of oedema produced by indomethacin, the effect of nitrogen mustard was not reversed by exogenous prostaglandin. In *b*, following incubation of plasma with zymosan for periods up to 40 min, zymosan was removed by centrifugation and the plasma was mixed with PGE_2 (3×10^{-10} mol per site) before injection into the skin. Exudation was measured over a 30-min period. Permeability-increasing activity was rapidly generated in plasma during incubation (response $t_{1/2}=100$ s) but oedema was only demonstrable when PGE_2 was present. Little exudation was induced by the following: plasma incubated alone, tested with PGE_2 (open symbols); plasma incubated with zymosan for 40 min but tested without PGE_2 (P* no E_2); or PGE_2 alone. Preheating plasma to 56 °C for 40 min (P* 56°40′) before the addition of zymosan prevented the generation of activity (40 min incubation with zymosan and the plasma tested with PGE_2). All the results are shown as the mean ±s.e.m. for six replicate injections. Values for PBS controls are shown as dashed lines (as in all the figures).

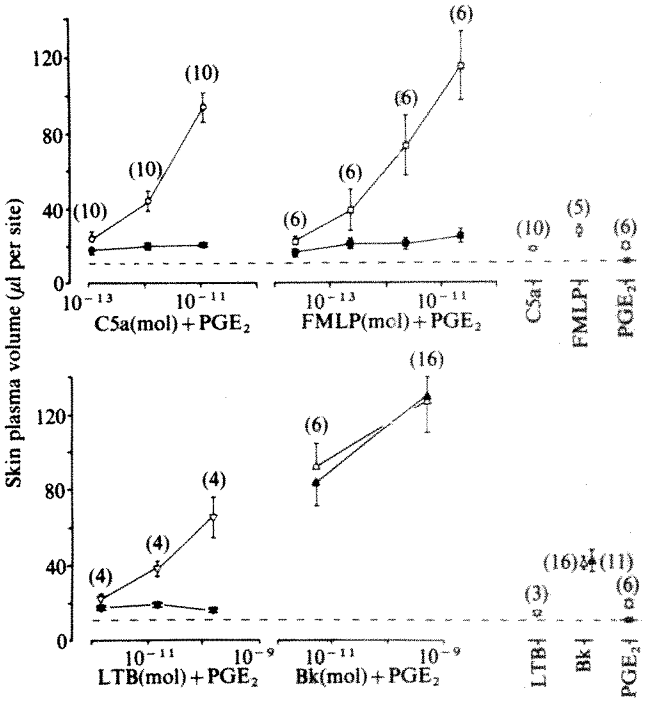


Fig. 2 Oedema induced in rabbit skin by rabbit C5a, FMLP, leukotriene B_4 (LTB) and bradykinin (Bk); the potentiating effect of PGE_2; and the inhibitory effect of leukocyte depletion. All responses were measured over a 30-min period. Numbers in parentheses refer to the number of animals used in each test; the mean response to six replicate injections of each substance was determined for each animal. Open symbols show responses in untreated animals. Closed symbols show responses in animals treated 4 days previously with a single intravenous injection of nitrogen mustard (1.75 mg per kg). The symbols on the right-hand side show responses to C5a, FMLP, leukotriene B_4 and bradykinin at the highest doses used in the experiments. PGE_2 was used at a dose of 3×10^{-10} mol per site throughout. In normal rabbits (PMN leukocyte count $3.5\pm1.1\times10^6$ per ml blood, $n=6$ rabbits) very small responses were elicited by C5a, FMLP, leukotriene B_4 and PGE_2 when tested alone (20–30 μl responses to LTB_4 alone at higher doses have been reported[30]). However, when mixed with PGE_2 the leukotactic substances were very potent in inducing oedema. Bradykinin produced distinct responses alone and these were enhanced by PGE_2. In PMN leukocyte-depleted rabbits (PMN leukocyte count $4.0\pm1.4\times10^4$ per ml blood, $n=6$ rabbits), although responses to bradykinin and bradykinin+PGE_2 were unaffected, responses to the leukotactic substances in combination with PGE_2 were virtually abolished. Responses to histamine+PGE_2, like those to bradykinin+PGE_2, were unaffected by PMN-leukocyte depletion—histamine (2.3×10^{-8} mol per site)+PGE_2 (3×10^{-10} mol per site) produced a response of 82.7±2.8 μl in controls and 78.0±14.5 μl in treated animals ($n=4$ rabbits).

Prostaglandins and oedema formation

The presence of microorganisms in a tissue stimulates leakage of plasma from small blood vessels resulting in tissue oedema. We investigated the mechanisms underlying this vascular response by injecting yeast cell walls (zymosan) into rabbit dorsal skin and measuring plasma leakage using the local accumulation of intravenously injected ^{131}I-albumin[6]. Zymosan induces acute oedema with a peak rate of plasma exudation at 1–3 h, decreasing to low levels by 4–5 h (ref. 7). Figure 1*a* shows that the volume of oedema fluid accumulating is regulated according to the stimulus size (the quantity of zymosan injected).

We deduced that zymosan stimulates the release in the skin of two chemicals which exert a synergistic action to mediate oedema ('the two mediator hypothesis'[8]). The first mediator increases vascular permeability by an action involving vascular endothelial cells; the second greatly potentiates plasma exudation by inducing relaxation of arteriolar smooth muscle cells and thus increasing the blood supply to the tissue. The vasodilator mediator is thought to be a prostaglandin[7–10] (PGE_2 or PGI_2)[9] produced by cell membranes; little exudation is induced when prostaglandin synthesis is inhibited (for example, by indomethacin[11]) as shown in Fig. 1*a*.

Involvement of the complement system

We next investigated the identity of the endogenous permeability-increasing mediator. We obtained evidence[12-14] that this mediator is derived from the complement system—the enzyme system in blood plasma and tissue fluid which is responsible for lysis of microorganisms. As shown in Fig. 1*b*, incubation of blood plasma with zymosan at 37°C resulted in the rapid generation of a substance which could increase vascular permeability when injected into skin. This substance induced little oedema unless mixed with a vasodilator prostaglandin. The active substance has been purified and identified[14] as a polypeptide fragment of the fifth component of complement, C5a. The activity is stable in plasma and is therefore probably due to C5a devoid of the carboxyl-terminal arginine (C5a des Arg); the arginine group, which is necessary for histamine-releasing (anaphylatoxic) activity in other species, is cleaved by plasma carboxypeptidase B (ref. 15). We have shown that oedema induced by synergism between rabbit C5a and prostaglandin is not dependent on histamine release[12,14].

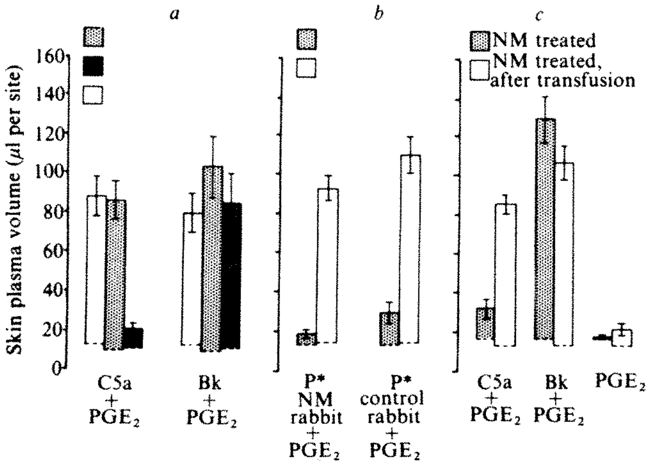


Fig. 3 *a*, Results of an experiment designed to show that nitrogen mustard suppresses responses in skin by an action on dividing bone marrow cells, rather than by an action on the skin itself. Rabbits were set up in three groups of four individuals. The first group were untreated (open columns). The second group were injected, via an ear vein, with nitrogen mustard (1.75 mg per kg) while a clamp was applied to the terminal descending aorta (stippled columns) and the clamp was removed after 10 min. As nitrogen mustard has a very short half life in the circulation, this procedure allowed access of nitrogen mustard to the dorsal skin but not to the bone marrow of the hind limbs. The third group were sham-operated and injected with nitrogen mustard via an ear vein (hatched columns). As in the second group a volume of blood (17% of total blood volume, measured by isotope dilution) was not exposed to nitrogen mustard; this volume of blood was withdrawn from the individuals of the sham-operated group before drug injection, maintained at 37 °C and reinjected after 10 min. Four days later, agents were injected into the skin and responses measured over a 30-min period. All three groups responded to bradykinin (5×10^{-10} mol per site) + PGE_2 (3×10^{-10} mol per site). Animals whose skin had been exposed to nitrogen mustard but whose bone marrow had been protected (group 2) responded normally to C5a (1.5×10^{-11} mol per site) + PGE_2 (3×10^{-10} mol per site) showing that inhibition of responses by nitrogen mustard was not due to an effect of the drug on the skin. *b*, Results of an experiment designed to investigate the effect of nitrogen mustard (NM) on C5a generation in plasma. Rabbits were set up in two groups of five individuals, one group injected with nitrogen mustard (stippled), the other untreated (open columns). Four days later plasma from each animal was incubated with zymosan (1 mg ml^{-1}, 37 °C for 30 min). The zymosan was then removed by centrifugation. Plasma was mixed with PGE_2 (3×10^{-10} mol per site) and responses measured in the skin (30 min period) of the donor animal and in an animal in the other group. Treated animals did not respond to plasma from either group; however, untreated animals responded to plasma from both groups. These results demonstrate that nitrogen mustard pretreatment suppresses responses to C5a, but not its production. *c*, Results of an experiment to show skin responses in a depleted animal before, and after, cross blood transfusion with a normal animal. Six such experiments were carried out; three depleted animals showed only a small elevation in circulating PMN leukocytes following transfusion and these animals did not respond to C5a + PGE_2. In three other experiments the circulating PMN leukocyte count was elevated following transfusion and these recipient animals responded to C5a + PGE_2. ^{125}I-albumin was injected intravenously into both normal and depleted animals. C5a (1.5×10^{-11} mol per site) + PGE_2 (3×10^{-10} mol per site) and bradykinin (5×10^{-10} mol per site) + PGE_2 (3×10^{-10} mol per site) were then injected intradermally into the depleted animal. After 30 min, blood was transfused (6 ml min^{-1}) by a peristaltic pump interposed between carotid artery cannulae and ear vein cannulae. After a further 60 min, pumping was stopped, the treated animal was injected intravenously with ^{131}I-albumin, and the skin was again tested with the same combinations of agents. Following transfusion the circulating PMN leukocyte count was elevated from 3×10^4 ml^{-1} to 2.4×10^6 ml^{-1}; only a small change was seen in the mononuclear leukocyte count, from 2.3×10^6 to 3.2×10^6 ml^{-1}. The animal clearly responded to C5a + PGE_2 following transfusion. Results shown as mean ± s.e.m. for six replicate injections. Hatched bars, nitrogen mustard-treated; open bars, nitrogen mustard-treated after transfusion.

Central role for the PMN leukocyte

In both intact and des Arg forms, C5a shows powerful chemoattraction for PMN leukocytes *in vitro*[16]. Thus, is the permeability-increasing activity of C5a in the skin dependent on the presence of PMN leukocytes *in vivo*? Zymosan failed to induce oedema in the skin of rabbits depleted of circulating PMN leukocytes using nitrogen mustard (results in Fig. 1 legend), providing the first indication that the action of C5a might be leukocyte dependent.

To investigate this further, the activity of histamine and bradykinin (both with established permeability-increasing activity) was compared with that of three leukotactic substances[16-18] in the skin of normal animals and animals depleted of circulating PMN leukocytes. The leukotactic substances used were: rabbit C5a (ref. 14), a polypeptide of approximately 70 amino acids; leukotriene B_4, a dihydroxy fatty acid formed from arachidonic acid by the action of leukocyte lipoxygenase[19]; and *N*-formyl-methionyl-leucyl-phenylalanine (FMLP), a synthetic substance based on formyl peptides secreted by bacteria in culture[18]. Figure 2 shows that in normal animals leukotactic substances alone induced little oedema, but, mixed with PGE_2, all three were highly potent permeability-increasing substances. However, in animals given intravenous nitrogen mustard 4 days before testing, to deplete circulating PMN leukocytes[20,21], the leukotactic substances were ineffective, even when combined with PGE_2. In normal animals both bradykinin and histamine induced some oedema when injected alone into the skin and the responses were greatly enhanced by PGE_2. However, these responses were unaffected by nitrogen mustard pretreatment. These results suggested that bradykinin and histamine increase vascular permeability by acting directly on vascular endothelial cells ('direct-action mediators'), but that C5a, leukotriene B_4 and FMLP can only act in the presence of PMN leukocytes ('polymorph-dependent mediators').

Nitrogen mustard (administered as a single intravenous injection) is short-lived in the blood and depletes circulating PMN leukocytes by preventing replication of stem cells in bone marrow. However, as the skin itself is exposed to the drug, a direct effect on the skin could not be excluded. A clamp applied to the terminal descending aorta during, and for 10 min after, an ear-vein injection of nitrogen mustard provided access for the drug to the skin, but not to the bone marrow of the hind limbs. As Fig. 3*a* shows, 4 days later animals treated in this way responded normally to C5a + PGE_2. Thus, suppression of responses by nitrogen mustard was not due to an effect on the skin. Furthermore, 4 days after nitrogen mustard treatment, the numbers of circulating mononuclear cells and platelets were little affected[20] and, as shown in Fig. 3*b*, the plasma of animals pretreated with nitrogen mustard could be activated normally to produce permeability-increasing activity, that is, C5a (complement levels measured by haemolytic assay are also known to be unaffected[20]).

All these results suggested that PMN leukocytes are essential for the permeability response to the leukotactic substances. Nevertheless, we could not raise circulating cell numbers sufficiently to restore responsiveness in previously depleted animals by intravenous or intra-arterial injections of PMN leukocytes. This probably indicates the importance of the surface properties of the cells which may change following removal from the body. Experiments using whole blood were more successful: in three out of six experiments it was possible to elevate circulating levels of PMN leukocytes in depleted rabbits by whole-blood cross-transfusion (6 ml min^{-1} for 60 min) using normal donor rabbits. Figure 3*c* shows that responsiveness to C5a was restored in these rabbits.

In all experiments Evans blue dye was mixed with ^{131}I-albumin to indicate plasma leakage sites in the skin. Responses to different agents were visually similar; however, direct-action mediators could be clearly distinguished from polymorph-dependent mediators by their time courses of action using radioisotope techniques, as shown in Fig. 4. Oedema was first detectable at 1.5 min after injection of bradykinin + PGE_2 but at 6 min with C5a + PGE_2 (surprisingly early considering the necessity for PMN leukocytes). A more striking difference was that responses to bradykinin and histamine were of very short duration, as expected from earlier experiments[22] (for example, bradykinin $t_{1/2} = 4$ min), whereas those to leukotactic substances were protracted (for example, C5a $t_{1/2} = 95$ min).

From our results we conclude that mechanisms exist which regulate local fluid accumulation in response to the presence of microorganisms in tissue. This process, although remarkably fast in onset, seems to be dependent on PMN leukocytes. PMN leukocytes are attracted to the site and their function is controlled by chemicals released in the tissue. The chemical signals can be generated in response to microorganisms in tissue fluid (for example, C5a), can be derived from microorganisms themselves (for example, FMLP), or, perhaps in later phases of the response, derived from leukocytes (for example, leukotriene B_4). These substances induce little oedema alone; the concomitant generation of vasodilator prostaglandin seems to be essential for oedema formation.

How do PMN leukocytes increase vascular permeability?

The manner in which PMN leukocytes increase vessel wall permeability remains undetermined. PMN leukocytes, during their passage through the vessel wall, may secrete enzymes which break down substances present at junctions between adjacent endothelial cells, or the enzymes may act on basement membrane underlying endothelial cells. Alternatively, PMN leukocytes may secrete a substance which affects endothelial cells, perhaps causing contraction. If PMN leukocytes do secrete such a substance, phagocytosis does not seem to be essential because soluble factors (C5a, FMLP and leukotriene B_4) or particulate matter (zymosan) can induce increased vascular permeability equally well.

Secretion by PMN leukocytes may be unnecessary as electron micrographs of tissues removed shortly after injection of C5a showed PMN leukocytes between the endothelial cell layer and basement membrane, many underneath endothelial cell junctions. Direct contact of that sort may be required for endothelial cells to respond. If PMN leukocytes located under junctions were to swell (as has been observed in response to C5a *in vitro*[23]) this could open the junctions.

PMN leukocytes in other types of inflammatory reaction?

We believe that these mechanisms have developed as an essential part of the host defence system to control the supply of complement from the blood to an infected tissue in order to regulate local killing of microorganisms. Similar mechanisms probably operate in allergic inflammation to regulate the supply of complement, antibodies and other plasma components to the tissue. In many models of inflammation oedema is known to be suppressed by prostaglandin synthesis inhibitors, complement depletion and PMN leukocyte depletion; these observations are consistent with our hypothesis.

Pulmonary oedema in man is often associated with a local accumulation of PMN leukocytes, for example in adult respiratory distress syndrome due to septicaemia, during haemodialysis and cardiopulmonary bypass[24]. There is evidence that the interaction of C5a with PMN leukocytes is important in these conditions[24,25].

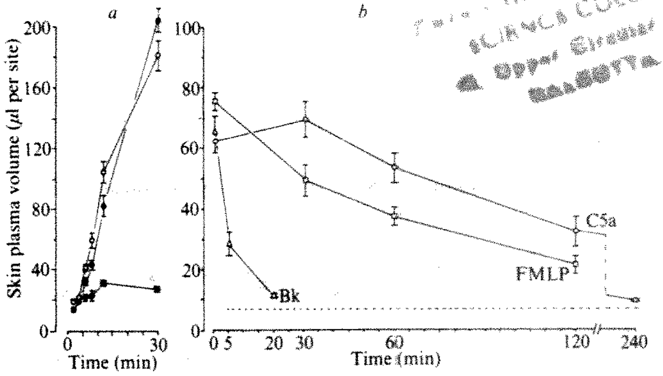


Fig. 4 The onset time of the vascular permeability increase induced by injection of C5a (*a*), and the duration of increased permeability induced by C5a and FMLP (*b*). *a*, Rabbits were injected intravenously with ^{131}I-albumin and killed 30 min later. Intradermal injections were given at intervals (shown in the abscissa) before death. The injections given were: ○, C5a (1.5×10^{-11} mol per site) + PGE_2 (3×10^{-10} mol per site); ●, C5a + PGE_2 + mepyramine maleate (3.5×10^{-9} mol per site); and ◆, PGE_2 alone. Six minutes after injection, responses to C5a + PGE_2 were significantly greater than control values produced by PGE_2 alone ($P < 0.01$, Student's unpaired *t*-test). Mepyramine, at doses which abolished matching responses to histamine + PGE_2, had a limited suppressor effect on responses; responses to C5a + PGE_2 + mepyramine were also significantly elevated above controls at 6 min ($P < 0.01$). Bradykinin and histamine responses were earlier in onset, for example bradykinin (5×10^{-10} mol per site) + PGE_2 (3×10^{-10} mol per site) onset time 1.5 min ($P < 0.01$). In the experiments shown in *b*, bradykinin (5×10^{-10} mol per site), C5a (1.5×10^{-11} mol per site) and FMLP (2.3×10^{-11} mol per site) were injected at intervals (as shown in abscissa) before a local superimposed injection of PGE_2 (3×10^{-10} mol per site). Labelled albumin was injected intravenously immediately before the PGE_2 injection and responses measured over a period of 30 min. (This procedure is designed to negate the effect of PGE_2 metabolism on the result.) Clearly the durations of action of the leukotactic substances were quite different from that of bradykinin: bradykinin $t_{1/2} = 4$ min, C5a $t_{1/2} = 95$ min, and FMLP $t_{1/2} = 50$ min. Histamine responses were similar in duration to those of bradykinin (results not shown). Results are represented as mean ± s.e.m., $n = 4$ replicates in *a*, and $n = 6$ replicates in *b*.

In chronic, self-destructive inflammation (such as rheumatoid arthritis), mononuclear cells are intimately involved with the underlying pathogenesis, and these cells could interact with the vessel wall in the manner we have described. In such chronic conditions, PMN leukocytes may be important because phases of active oedema formation are associated with the appearance of high numbers of these cells in the tissues[26]. Our simple model may be relevant to these complex events as rheumatoid fluid has been shown to contain prostaglandins[27], leukotriene B_4 (ref. 28) and C5a (ref. 29).

In summary, we suggest that PMN leukocytes and endothelial cells lining the blood vessel wall cooperate to regulate the supply of macromolecules to a tissue, and that this may represent a fundamental component of the inflammatory response.

We thank our colleagues in the Department of Pharmacology, Royal College of Surgeons, for discussions. C5a, from zymosan-activated rabbit plasma, was purified and characterized in collaboration with Dr P. J. Jose[14]. We thank Dr A. W. Ford-

Hutchinson and his colleagues for the gift of leukotriene B_4, and M. J. Peck and M. J. Forrest for technical assistance. This work was supported by the MRC.

Received 12 August; accepted 2 December 1980.

1. Winter, C. A., Risley, E. A. & Nuss, G. V. *Proc. Soc. exp. Biol. Med.* **111,** 544 (1962).
2. Di Rosa, M., Giroud, J. P. & Willoughby, D. A. *J. Path. Bact.* **104,** 15–29 (1971).
3. Arthus, M. *C.r. Séanc. Soc. Biol.* **55,** 817 (1903).
4. Metchnikoff, E. *Annls Inst. Pasteur, Paris* **1,** 321 (1887).
5. Movat, H. Z. *Exp. molec. Path.* **31,** 201–210 (1979).
6. Williams, T. J. *J. Physiol., Lond.* **254,** 4-5P (1976).
7. Williams, T. J. *Br. J. Pharmac.* **65,** 517–524 (1979).
8. Williams, T. J. & Peck, M. J. *Nature* **270,** 530–532 (1977).
9. Peck, M. J. & Williams, T. J. *Br. J. Pharmac.* **62,** 464–465P (1978).
10. Williams, T. J. & Morley, J. *Nature* **246,** 215–218 (1973).
11. Vane, J. R. *Nature new Biol.* **231,** 232–235 (1971).
12. Williams, T. J. *J. Physiol., Lond.* **281,** 44–45P (1978).
13. Jose, P. J., Peck, M. J., Robinson, C. & Williams, T. J. *J. Physiol., Lond.* **281,** 13–14P (1978).
14. Williams, T. J. & Jose, P. J. *J. exp. Med.* **153,** 136–153 (1981).
15. Hugli, T. E. & Müller-Eberhard, H. J. *Adv. Immun.* **26,** 1–53 (1978).
16. Fernandez, H. N., Henson, P. M., Otani, A. & Hugli, T. E. *J. Immun.* **120,** 109–115 (1978).
17. Ford-Hutchinson, A. W., Bray, M. A., Doig, M. V., Shipley, M. E. & Smith, M. J. H. *Nature* **286,** 264–265 (1980).

Note added in proof: The synergistic effect of leukotriene B_4 and PGE_2 has been confirmed by G. A. Higgs (personal communication) and in several species by M. A. Bray *et al.*[31].

18. Schiffmann, E., Corcoran, B. A. & Wahl, S. M. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1059–1062 (1975).
19. Samuelsson, B., Borgeat, P., Hammarström, S. & Murphy, R. C. *Prostaglandins* **17,** 785–787 (1979).
20. Humphrey, J. H. *Br. J. exp. Path.* **36,** 268–282 (1955).
21. Stetson, C.A. *J. exp. Med.* **94,** 347–348 (1951).
22. Williams, T. J. & Morley, J. *Br. J. exp. Path.* **55,** 1–12 (1974).
23. O'Flaherty, J. T., Kreutzer, D. L. & Ward, P. A. *Am. J. Path.* **90,** 537–550 (1978).
24. Larsen, G. L., McCarthy, K., Webster, R. O., Henson, J. & Henson, P. M. *Am. J. Path.* **100,** 179–192 (1980).
25. Hammerschmidt, D. E., Weaver, L. J., Hudson, L. D., Craddock, P. R. & Jacob, H. S. *Lancet* **i,** 947–949 (1980).
26. Ward, P. A. & Zvaifler, N. J. *J. clin. Invest.* **50,** 606–616 (1971).
27. Higgs, G. A., Vane, J. R., Hart, F. D. & Wotjulewski, J. A. in *Prostaglandin Synthetase Inhibitors* (eds Robinson, H. J. & Vane, J. R.) 165–174 (Raven, New York, 1974).
28. Klickstein, L. B., Shapleigh, C. & Goetzl, E. J. *J. clin. Invest.* **66,** 1166–1170 (1980).
29. Zvaifler, N. J. *J. exp. Med.* **134,** 276–285 (1971).
30. Eakins, K. E., Higgs, G. A., Moncada, S., Salmon, J. A. & Spayne, J.A. *J. Physiol., Lond.* **307,** 71P (1980).
31. Bray, M. A., Cunningham, F. M., Ford-Hutchinson, A. W. & Smith, M. J. H. *Br. J. Pharmac.* (in the press).

Nucleotide and corresponding amino acid sequences encoded by α and β tubulin mRNAs

P. Valenzuela, M. Quiroga, J. Zaldivar, W. J. Rutter, M. W. Kirschner & D. W. Cleveland

Department of Biochemistry and Biophysics, School of Medicine, University of California, San Francisco, California 94143

Most of the mRNA sequences coding for α and β tubulin in embryonic chick brain have been determined by sequencing of cloned cDNA copies of these mRNAs. From a 1,682-base pair cDNA sequence we have deduced the entire protein sequence for β tubulin. For α tubulin, all but about 38 N-terminal amino acids have been deduced from the cDNA sequence. Although tyrosine has previously been shown to be post-translationally added to the C-terminus of α tubulin by a specific ligase, we conclude that the primary post-translational event must be the removal, not the addition, of tyrosine because a terminal tyrosine is encoded by the mRNA.

TUBULIN is the principle subunit of microtubules, an intracellular cylindrical filamentous structure present in almost all eukaryotic cells. Microtubules function as structural and motile elements in mitosis, intracellular transport, flagellar movement and the cytoskeleton (for review see refs 1, 2). The microtubule, 24 nm in diameter, is comprised of 13 protofilaments each containing an alternating array of α and β tubulin subunits spaced every 4 nm. In the soluble form it is generally assumed that tubulin is a heterodimer composed of one α and one β chain each of molecular weight (MW) about 55,000.

The structure of the microtubule is highly conserved and at the resolution obtained by electron microscopy (~2 nm) no differences in structure have been observed between microtubules from *Tetrahymena*, sea urchin and mammalian sources. The tubulin subunit will form copolymers across wide species lines and the conditions for assembly of tubulin are similar for each species. Recently, we have shown that cDNA clones for chicken embryonic α and β tubulin (which do not cross-hybridize to each other) will hybridize in stringent conditions to genomic DNA from sea urchin and *Drosophila*, as well as mammals[3,4], and in less stringent conditions to yeast, *Chlamydomonas* and trypanosomes (N. Neff, D. Botstein, D. Cleveland and N. Agabian, unpublished results).

This high degree of conservation is not surprising in view of the large number of conserved structural interactions of tubulin. The tubulin dimer has two GTP binding sites, one exchangeable and the other not, a single colchicine binding site, two binding sites for vinblastine, and other drug binding sites. In addition, the tubulin structure must be constrained by the requirement for interaction with other tubulin molecules in the microtubule lattice as well as with microtubule-associated proteins.

Despite considerable recent interest in the biochemistry of tubulin, there are important gaps in our knowledge of its primary structure and genetic complexity. Evidence suggesting some homology and possible common evolutionary origin of the α and β peptides relies on the published N-terminal 25 amino acids in α and β tubulin[5]. Additional other known data are the sequence of some unordered peptides[6] and the 78 C-terminal amino acids of the α chain[7]. A true appreciation of the degree of homology between α and β chains, however, awaits more complete sequence information. Additional sequencing studies are also crucial to understanding the microheterogeneity, post-translational modification and structure of the tubulin proteins themselves.

Recently, we have constructed cDNA clones from embryonic mRNA of chicken brain and have isolated clones containing α or β tubulin sequences. We report here the complete DNA sequence of the cDNA from these clones. The β tubulin clone

0028–0836/81/080650—06$01.00

contains the entire coding sequence for the protein and extensive 5′ and 3′ untranslated regions of the mRNA. The α clone contains >90% of the coding sequences and some of the 3′ untranslated region of the α tubulin mRNA.

DNA sequences of plasmids pT2 and pT1

The construction of bacterial plasmids containing sequences homologous to α or β tubulin mRNAs from embryonic chick brain has been described elsewhere[3]. We initially demonstrated by mRNA hybridization followed by *in vitro* translation of specifically bound mRNA that the hybrid plasmids pT2 and pT1 carry β and α tubulin sequences, respectively. The sequencing strategies for both pT1 and pT2 are presented in Fig. 1. All sequencing was done by the method of Maxam and Gilbert[8]. In all cases both strands of the cDNAs were sequenced unless a particular single-strand sequence was exceptionally clear.

The sequence of the 1,681 nucleotides of the cDNA from pT2, the β tubulin plasmid, is given in Fig. 2*a*, along with the corresponding protein sequence which it predicts. The direction and frame of reading were established by comparison with the published N-terminal 25-amino acid sequence for chick brain β tubulin[5] and by the presence of the 3′ poly(A) tract at one end of the cDNA. In addition, the C-terminus of the protein was established by the appearance of three in-phase stop codons and by the presence of a C-terminal peptide with a composition (15 out of 25 residues acidic) similar to that reported for α tubulin[7]. Plasmid pT2 thus contains the complete coding sequence for the 445 amino acids of β tubulin as well as 87 nucleotides 5′ to the methionine initiation codon and 226 nucleotides between the UGA stop codon and the site of polyadenylation. The predicted MW of β tubulin is 49,935 and the sequence deduced agrees with 19 of the 25 N-terminal residues reported by Luduena and Woodward[5] for chick brain β tubulin. The cDNA sequence predicts that residues 11 and 15 are glutamic acid, whereas glutamine was the amino acid residue previously reported[5] at these positions. Similarly, we predict that residues 10, 13 and 17 are glycine whereas threonine was originally reported at these positions.

The sequence of the chick brain α tubulin clone, pT1, is given in Fig. 2*b*. The direction and frame of reading were established by identification of the known C-terminal protein sequences for porcine brain α tubulin. Although pT1 contains 123 nucleotides 3′ to the UGA stop codon, no poly(A) tract is present; therefore some regions of the 3′ untranslated region of the mRNA may be missing. Furthermore, pT1 does not contain the complete 5′

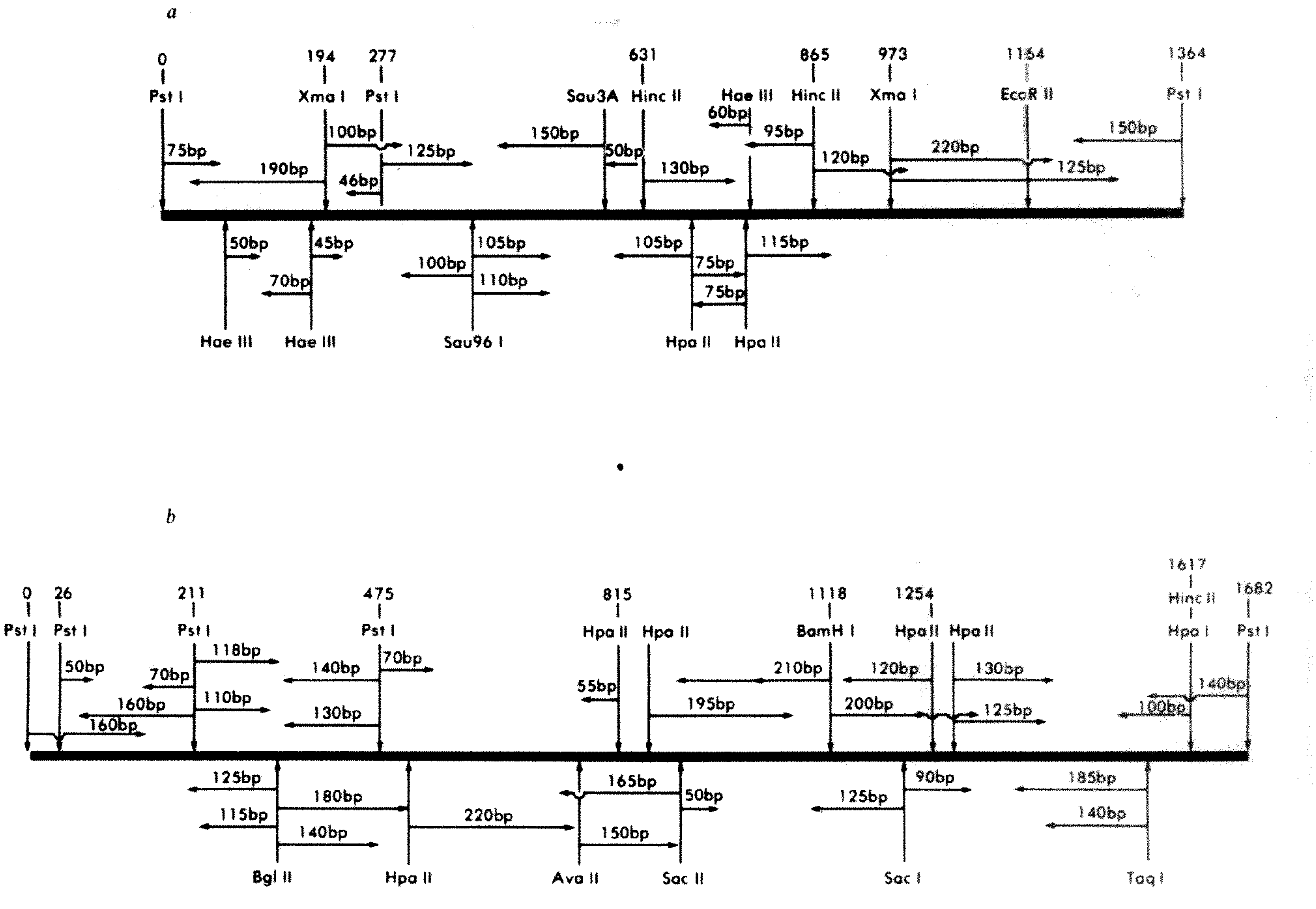


Fig. 1 Sequencing strategy for the chicken brain α and β tubulin cDNAs. *a*, α tubulin cDNA; *b*, β tubulin cDNA. Nucleotide residues are numbered in the direction from 5′ to 3′ in the message strand. The extent and direction of each fragment sequenced is shown by the horizontal arrows. Only the restriction sites used as starting points are shown. For end labelling, DNA was treated with 5 μg of bacterial alkaline phosphatase (Worthington) at 37 °C for 60 min in 20 mM Tris-HCl, *p*H 8.0, phenol-extracted and treated with 10 units of T4 polynucleotide kinase (Boehringer) at 37 °C for 30 min in a reaction mixture containing 50 mM glycine, *p*H 9.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.1 mM spermidine and 0.001 mM [γ-^{32}P]ATP (~3,500 Ci $mmol^{-1}$). Fragments labelled at only one end were isolated by gel electrophoresis after either digestion with a restriction enzyme or strand separation. DNA sequencing was carried out by the method of Maxam and Gilbert[8].

a

```
                                                                                        Met Arg Glu Ile Val His Ile Gln
UAAGGCCCGGAGCGGGGCGCGGCCGCUGCAGCGGAGAGCGGGAGGUGACGGAGCGGGAGCAGCGCGGCACCGGCAGACACCGGCAUC AUG CGU GAG AUC GUG CAC AUC CAG

    10                                      20                                      30
Ala Gly Gln Cys Gly Asn Gln Ile Gly Ala Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Asp Pro Thr Gly Ser Tyr His Gly
GCC GGC CAG UGC GGC AAC CAG AUC GGC GCU AAG UUC UGG GAG GUC AUC AGC GAU GAG CAC GGC AUC GAU CCC ACC GGC AGC UAC CAC GGG

    40                                      50                                      60
Asp Ser Asp Leu Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala Thr Gly Asn Lys Tyr Val Pro Arg Ala Ile Leu Val Asp Leu
GAC AGC GAC CUG CAG CUG GAG AGG AUC AAC GUG UAC UAC AAU GAG GCC ACA GGU AAC AAG UAC GUC CCC CGU GCC AUC CUG GUU GAC CUG

    70                                      80                                      90
Glu Pro Gly Thr Met Asp Ser Val Arg Ser Gly Pro Phe Gly Gln Ile Phe Arg Pro Asp Asn Phe Val Phe Gly Gln Ser Gly Ala Gly
GAG CCC GGC ACG AUG GAC UCG GUG CGC UCC GGC CCC UUU GGA CAG AUC UUC CGA CCC GAC AAC UUU GUC UUU GGU CAG AGC GGG GCC GGC

    100                                     110                                     120
Asn Asn Trp Ala Lys Gly His Tyr Thr Glu Gly Ala Glu Leu Val Asp Ser Val Leu Asp Val Val Arg Lys Glu Ser Glu Ser Cys Asp
AAC AAC UGG GCC AAG GGG CAC UAC ACG GAA GGC GCU GAG CUG GUG GAC UCU GUG CUG GAU GUG GUG AGG AAG GAG UCG GAA AGC UGC GAC

    130                                     140                                     150
Cys Leu Gln Gly Phe Gln Leu Thr His Ser Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu
UGC CUG CAG GGC UUC CAG UUG ACC CAC UCG CUG GGC GGC GGC ACC GGC UCG GGG AUG GGC ACC CUC CUC AUC AGC AAG AUC CGC GAG GAG

    160                                     170                                     180
Tyr Pro Asp Arg Ile Met Asn Thr Phe Ser Val Met Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser
UAC CCC GAC CGC AUC AUG AAC ACG UUC AGC GUC AUG CCC UCC CCC AAG GUG UCG GAC ACG GUG GUG GAG CCC UAC AAU GCC ACC CUU UCU

    190                                     200                                     210
Val His Gln Leu Val Glu Asn Thr Asp Glu Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr
GUG CAC CAG CUG GUG GAG AAC ACG GAC GAG ACC UAC UGC AUC GAC AAC GAG GCC CUG UAC GAC AUC UGC UUC CGC ACC CUG AAG CUC ACC

    220                                     230                                     240
Thr Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser Ala Thr Met Ser Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala
ACU CCC ACG UAC GGG GAC CUC AAC CAC CUG GUG UCG GCC ACC AUG AGC GGC GUG ACC ACC UGC CUU CGC UUC CCC GGC CAG CUG AAC GCC

    250                                     260                                     270
Asp Leu Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg Leu His Phe Phe Met Pro Gly Phe Ala Pro Leu Thr Ser Arg Gly Ser
GAC CUG CGC AAG CUG GCG GUC AAC AUG GUG CCU UUC CCC CGG CUG CAC UUC UUC AUG CCG GGC UUC GCC CCG CUG ACG AGC CGC GGC AGC

    280                                     290                                     300
Gln Gln Tyr Arg Ala Leu Thr Val Pro Glu Leu Thr Gln Gln Met Phe Asp Ser Lys Asn Met Met Ala Ala Cys Asp Pro Arg His Gly
CAG CAG UAC CGA GCC CUG ACG GUG CCC GAG CUG ACG CAG CAG AUG UUC GAC UCC AAG AAC AUG AUG GCC GCC UGC GAC CCC CGC CAC GGC

    310                                     320                                     330
Arg Tyr Leu Thr Val Ala Ala Ile Phe Arg Gly Arg Met Ser Met Lys Glu Val Asp Glu Gln Met Leu Asn Val Gln Asn Lys Asn Ser
CGC TAC CUG ACG GUG GCU GCC AUC UUC CGA GGC CGC AUG UCC AUG AAG GAG GUG GAC GAG CAG AUG CUG AAC GUG CAG AAC AAG AAC AGC

    340                                     350                                     360
Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro Arg Gly Leu Lys Met Ser Ala Thr Phe Ile
AGC UAC UUU GUG GAG UGG AUC CCC AAC AAC GUG AAG ACG GCC GUC UGC GAC AUC CCC CCG CGU GGC CUC AAG AUG UCC GCC ACC UUC AUC

    370                                     380                                     390
Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr
GGC AAC AGC ACG GCC AUC CAG GAG CUC UUC AAG AGG AUC UCG GAG CAG UUC ACG GCC AUG UUC CGG CGC AAG GCC UUC UUG CAC UGG UAC

    400                                     410                                     420
Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn Met Asn Asp Leu Val Ser Glu Tyr Gln Gln Tyr Gln Asp Ala
ACC GGC GAG GGC AUG GAU GAG AUG GAG UUC ACC GAG GCG GAG AGC AAC AUG AAC GAC CUG GUC UCA GAG UAC CAG CAA UAC CAG GAU GCC

    430                                     440             445
Thr Ala Asp Glu Gln Gly Glu Phe Glu Glu Glu Gly Glu Glu Asp Glu Ala OP
ACC GCU GAU GAG CAG GGU GAA UUU CAA GAG GAA GGA GAG GAG GAU GAG GCA UGA AGUGGAAAUGGGUAGGAGUCAAGUGUAGUCUGAGCAGGCAAGUGUUCA

UUGAGAGGAGAAUCUGUGCAGUUGUGCUGAAGCAUGCGUUUCUUAAUUUGGUGCUCUCCGCUGUUGCUUUUGUCAGCAGUUUUGUAUCCUCGACUGUCCGAUGUAACAGUUGCAAAAAUA

CUUCAGAGUCUUCUGUUGAAAUGGUUAACUUCUCAAACAUAAAGGCUUUUUGUGUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

b

```
        40                                  50                                      60          ●
    Thr Ile Gly Gly Gly Asp Asp Ser Phe Asn Thr Phe Phe Ser Glu Thr Gly Ala Gly Lys His Val Pro Arg Ala Val Phe Val Asp
GAA ACC AUC GGC GGG GGG GAC GAC UCC UUC AAC ACC UUC UUC AGC GAG ACG GGG GCC GGC AAG CAC GUG CCC CGU GCC GUC UUU GUG GAC

    70                                      80                                      90
Leu Glu Pro Thr Val Ile Asp Glu Val Arg Thr Gly Thr Tyr Arg Gln Leu Phe His Pro Glu Gln Leu Ile Thr Gly Lys Glu Asp Ala
CUG GAG CCC ACC GUG AUC GAU GAG GUG CGC ACG GGG ACG UAC CGG CAG CUC UUC CAC CCC GAG CAG CUG AUC ACG GGC AAG GAG GAU GCG

    100                                     110                                     120
Ala Asn Asn Tyr Ala Arg Gly His Tyr Thr Ile Gly Lys Glu Ile Ile Asp Leu Val Leu Asp Arg Ile Arg Lys Leu Ala Asp Gln Cys
GCC AAC AAC UAC GCC CGG GGG CAC UAC ACC AUC GGC AAG GAG AUC AUC GAC CUG GUG CUG GAC CGC AUC CGC AAG CUG GCU GAC CAG UGC

    130                                     140                                     150
Thr Gly Leu Gln Gly Phe Ser Val Phe His Ser Phe Gly Gly Gly Thr Gly Ser Gly Phe Thr Ser Leu Leu Met Glu Arg Leu Ser Val
ACG GGG CUG CAG GGC UUC UCG GUG UUC CAC AGC UUC GGG GGC GGC ACC GGC UCC GGC UUC ACC UCG CUG CUG AUG GAG CGC CUG UCG GUG

    160                                     170                                     180
Asp Tyr Gly Lys Lys Ser Lys Leu Glu Phe Ser Ile Tyr Pro Ala Arg Gln Val Ser Thr Ala Val Val Glu Pro Tyr Asn Ser Ile Leu
GAC UAC GGC AAG AAG UCC AAG CUG GAG UUC UCC AUC UAC CCG GCC CGU CAG GUG UCC ACG GCC GUG GUG GAG CCC UAC AAC UCC AUC CUC

    190                                     200                                     210
Thr Thr His Thr Thr Leu Glu His Ser Asp Cys Ala Phe Met Val Asp Asn Glu Ala Ile Tyr Asp Ile Cys Arg Arg Asn Leu Asp Ile
ACC ACC CAC ACC ACG CUG GAG CAC UCC GAC UGC GCC UUC AUG GUG GAC AAC GAG GCC AUC UAC GAC AUC UGU CGC AGG AAC CUG GAC AUC

    220                                     230                                     240
Glu Arg Pro Thr Tyr Thr Asn Leu Asn Arg Leu Ile Gly Gln Ile Val Ser Ser Ile Thr Ala Ser Leu Arg Phe Asp Gly Ala Leu Asn
GAG CGC CCC ACC UAC ACC AAC CUC AAC AGG CUG AUA GGG CAG AUC GUG UCC UCC AUC ACA GCC UCA CUG CGC UUC GAU GGG GCU CUG AAC

    250                                     260                                     270
Val Asp Leu Thr Glu Phe Gln Thr Asn Leu Val Pro Tyr Pro Arg Ile His Phe Pro Leu Ala Thr Tyr Ala Pro Val Ile Ser Ala Glu
GUC GAC CUG ACG GAG UUC CAG ACC AAC CUG GUG CCC UAC CCU CGC AUC CAC UUC CCA UUG GCC ACG UAU GCG CCG GUG AUC UCU GCU GAG

    280                                     290                                     300
Lys Ala Tyr His Glu Gln Leu Ser Val Ala Glu Ile Thr Asn Ala Tyr Phe Glu Pro Ala Asn Gln Met Val Lys Cys Asp Pro Arg His
AAG GCC UAC CAC GAG CAG CUG UCG GUG GCU GAG AUC ACC AAC GCG UAC UUC GAG CCG GCC AAC CAG AUG GUG AAG UGC GAC CCG CGG CAC

    310                                     320                                     330
Gly Lys Tyr Met Ala Cys Cys Leu Leu Tyr Arg Gly Asp Val Val Pro Lys Asp Val Asn Ala Ala Ile Ala Thr Ile Lys Thr Lys Arg
GGC AAG UAC AUG GCG UGC UGC CUG CUG UAC CGC GGG GAC GUG GUG CCC AAA GAU GUC AAC GCC GCC AUC GCC ACC AUC AAG ACC AAG CGC

    340                                     350                                     360
Thr Ile Gln Phe Val Asp Trp Cys Pro Thr Gly Phe Lys Val Gly Ile Asn Tyr Gln Pro Pro Thr Val Val Pro Gly Gly Asp Leu Ala
ACC AUU CAG UUC GUA GAC UGG UGC CCA ACU GGU UUC AAG GUG GGC AUC AAC UAC CAG CCU CCC ACC GUG GUG CCC GGG GGG GAC CUG GCC

    370                                     380                                     390
Lys Val Gln Arg Ala Val Cys Met Leu Ser Asn Thr Thr Ala Ile Ala Glu Ala Trp Ala Arg Leu Asp His Lys Phe Asp Leu Met Tyr
AAG GUG CAG CGC GCC GUG UGC AUG CUG AGC AAC ACC ACG GCC AUC GCC GAG GCG UGG GCG CGC UUG GAC CAC AAG UUU GAC CUG AUG UAC

    400                                     410                                     420
Ala Lys Arg Ala Phe Val His Trp Tyr Val Gly Glu Gly Met Glu Glu Gly Glu Phe Ser Glu Ala Arg Glu Asp Met Ala Ala Leu Glu
GCC AAG CGC GCC UUC GUG CAC UGG UAU GUG GGG GAG GGC AUG GAG GAG GGG GAG UUC UCA GAG GCU CGG GAG GAU AUG GCU GCC CUG GAG

    430                                     440                                     450
Lys Asp Tyr Glu Glu Val Gly Val Asp Ser Val Glu Gly Glu Gly Glu Glu Glu Gly Glu Glu Tyr OP
AAG GAU UAC GAG GAG GUG GGG GUG GAU UCC GUG GAG GGU GAG GGU GAG GAG GAA GGA GAG GAG UAC UGA GAGUUCACAACGGUGCUGCUCGUACAGGGA

ACAAUCCUAGUUAUCCCCCACCCCACCCCCGUGGGUCUGUAGCCGUGUGCGCUGCUGUCUGUGCUGUCUGUGGUCUGGGUUCAAUGGAAAGUC
```

Fig. 2 *a*, The nucleotide sequence of chicken brain *β* tubulin mRNA, determined from the double-stranded cDNA insert of pT2 (ref. 3). The predicted complete amino acid sequence is shown above the RNA sequence. *b*, The partial nucleotide sequence of chicken brain *α* tubulin mRNA, determined from the double-stranded cDNA insert of pT1 (ref. 3). The predicted partial amino acid sequence of *α* tubulin is also indicated. The amino acids have been numbered by comparison with the complete sequence of *β* tubulin mRNA shown in *a*.

sequences of the mRNA. By alignment with the deduced β tubulin protein sequence from above, we predict that pT1 lacks the sequences coding for amino acids 1–38 of α tubulin. We used this assumption tentatively to number the amino acids for the α chain in Fig. 2*b*. The predicted MW for α tubulin is ~50,040. The predicted sequence corresponds to the existing protein sequence data in the following respects: (1) with the exception of the terminal tyrosine residue (see below), the C-terminal 78 amino acids reported for porcine brain α tubulin are identical to the chick sequence reported here, and (2) the sequences of four (α4D, α4HB, α4I and α4CA) of the eight unordered cyanogen bromide fragments reported for beef brain α tubulin[6] are identical to peptides in the chick sequence. One of the remaining reported peptides, α4J, contains a single substitution at position 302 in our sequence, and one peptide (α4HA) is actually a β tubulin peptide (residues 148–164).

a	2/UUU/Phe	1/UCU/Ser	2/UAU/Tyr	1/UGU/Cys
	18/UUC/Phe	9/UCC/Ser	17/UAC/Tyr	7/UGC/Cys
	0/UUA/Leu	2/UCA/Ser	0/UAA/OC	1/UGA/OP
	2/UUG/Leu	5/UCG/Ser	0/UAG/AM	3/UGG/Trp
	0/CUU/Leu	2/CCU/Pro	0/CAU/His	2/CGU/Arg
	3/CUC/Leu	9/CCC/Pro	11/CAC/His	13/CGC/Arg
	0/CUA/Leu	2/CCA/Pro	0/CAA/Gln	0/CGA/Arg
	24/CUG/Leu	4/CCG/Pro	12/CAG/Gln	4/CGG/Arg
	1/AUU/Ile	1/ACU/Thr	0/AAU/Asn	0/AGU/Ser
	21/AUC/Ile	18/ACC/Thr	15/AAC/Asn	3/AGC/Ser
	1/AUA/Ile	1/ACA/Thr	1/AAA/Lys	0/AGA/Arg
	8/AUG/Met	10/ACG/Thr	17/AAG/Lys	2/AGG/Arg
	0/GUU/Val	6/GCU/Ala	7/GAU/Asp	3/GGU/Gly
	3/GUC/Val	22/GCC/Ala	18/GAC/Asp	13/GGC/Gly
	1/GUA/Val	0/GCA/Ala	1/GAA/Glu	1/GGA/Gly
	28/GUG/Val	6/GCG/Ala	33/GAG/Glu	15/GGG/Gly

b	5/UUU/Phe	2/UCU/Ser	0/UAU/Tyr	0/UGU/Cys
	18/UUC/Phe	5/UCC/Ser	16/UAC/Tyr	8/UGC/Cys
	0/UUA/Leu	1/UCA/Ser	0/UAA/OC	1/UGA/OP
	2/UUG/Leu	7/UCG/Ser	0/UAG/AM	4/UGG/trp
	2/CUU/Leu	1/CCU/Pro	0/CAU/His	3/CGU/Arg
	6/CUC/Leu	16/CCC/Pro	10/CAC/His	11/CGC/Arg
	0/CUA/Leu	0/CCA/Pro	1/CAA/Gln	3/CGA/Arg
	22/CUG/Leu	3/CCG/Pro	21/CAG/Gln	2/CGG/Arg
	0/AUU/Ile	1/ACU/Thr	2/AAU/Asn	0/AGU/Ser
	19/AUC/Ile	15/ACC/Thr	21/AAC/Asn	14/AGC/Ser
	0/AUA/Ile	1/ACA/Thr	0/AAA/Lys	0/AGA/Arg
	19/AUG/Met	13/ACG/Thr	15/AAG/Lys	3/AGG/Arg
	1/GUU/Val	4/GCU/Ala	7/GAU/Asp	3/GGU/Gly
	7/GUC/Val	20/GCC/Ala	19/GAC/Asp	25/GGC/Gly
	0/GUA/Val	1/GCA/Ala	5/GAA/Glu	2/GGA/Gly
	21/GUG/Val	2/GCG/Ala	31/GAG/Glu	5/GGG/Gly

Fig. 4 Codon usage in α and β tubulin mRNAs. The numbers indicate the frequency with which individual codons are used in the coding regions of the mRNAs. OC, OP and AM designate the terminator codons ochre, opal and amber, respectively. *a*, α tubulin mRNA; *b*, β tubulin mRNA.

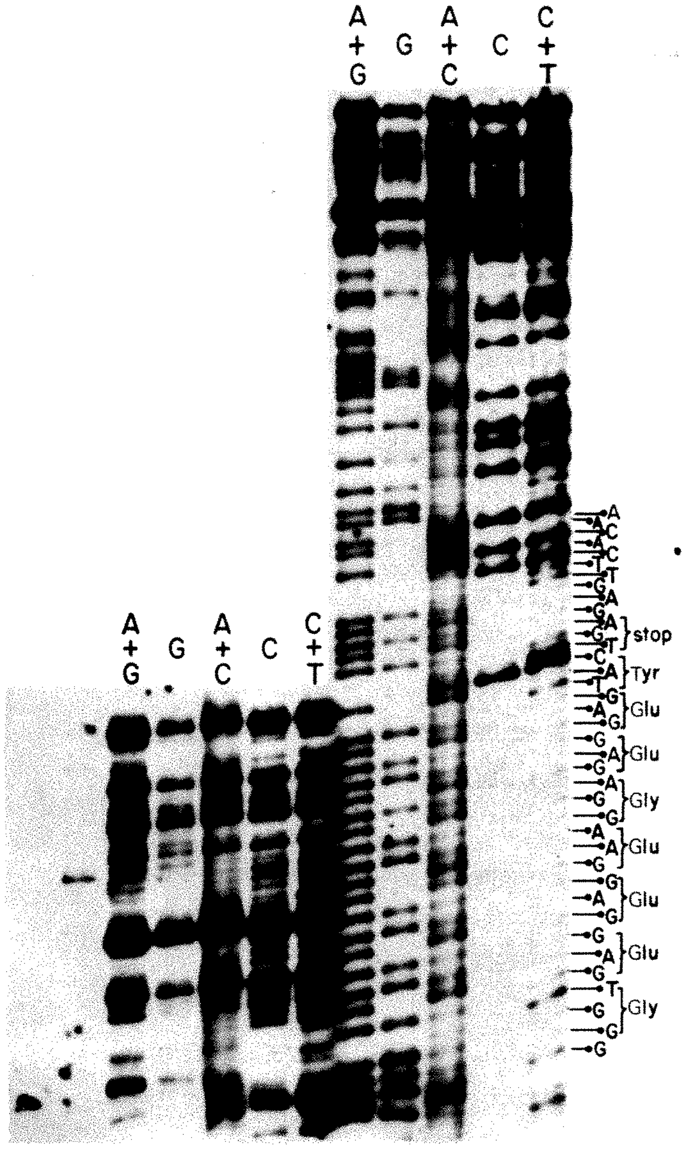


Fig. 3 Autoradiogram of a sequencing gel showing the region coding for the C-terminus of α tubulin. Note that the last codon before the UGA stop codon is UAC, which codes for tyrosine.

The tubulin subunits have been widely seen to migrate on SDS or SDS-urea polyacrylamide gels with mobilities characteristic of proteins of MW ~55,000–56,000 for α tubulin and 53,000–55,000 for β tubulin (for example, see ref. 9). Comparable mobilities have been found for the *in vitro* translation products of α and β tubulin from mRNA translated in reticulocyte or wheat-germ systems (for example, see ref. 10). Surprisingly, we have unambiguously determined that the polypeptide encoded by one β tubulin mRNA has a MW of 49,935. Similarly, although we cannot be completely certain of the size of the α tubulin chain, by homology with the β tubulin sequence and by comparison with unpublished protein sequence data of porcine brain α tubulin (H. Ponstingl and collaborators, unpublished results), we predict that the MW of the α tubulin chain is slightly greater than 50,000. Because the mobilities of *in vitro* synthesized chains are indistinguishable from those of the mature polypeptides, post-translational modification cannot account for the differences in estimated and actual molecular weights.

Note also that the β chain contains 19 methionines and the α chain probably contains 11 methionines (using the data of Luduena and Woodward[5] for amino acids 1–25 coupled with the sequence of cyanogen bromide fragment α4CB of Lu and Elzinga[6], which seems to be derived from amino acids 22–36 of the α chain of beef α tubulin). In agreement with previous results[5], this predicts a ratio in methionine contents of 1.7:1, that is, the expected ratio of incorporation of label into each chain if each are translated at comparable rates or in comparable amounts. This could be a useful number for investigators using methionine-labelled materials.

The C-terminal residue of *α* tubulin is tyrosine

An enzyme has recently been reported which is capable of post-translationally adding a tyrosine residue to the C-terminal end of *α* tubulin[11-14]. The enzyme is specific for *α* tubulin, requires ATP and functions in the presence of protein synthesis inhibitors. Purified *α* tubulin has been reported to contain between 15 and 50% tyrosine on the C-terminal end[7,8,15]. Historically, then, C-terminal tyrosylation has been viewed as a post-translational modification of the *α* chain. However, it is now clear from the nucleotide sequence that *α* tubulin mRNA contains a coded terminal tyrosine (Fig. 2*b*). The sequencing gel demonstrating the tyrosine UAC codon immediately before the UGA stop codon is shown in Fig. 3. Although the functional significance of tyrosylation and detyrosylation remains unclear, these new data demonstrate that the primary post-translational

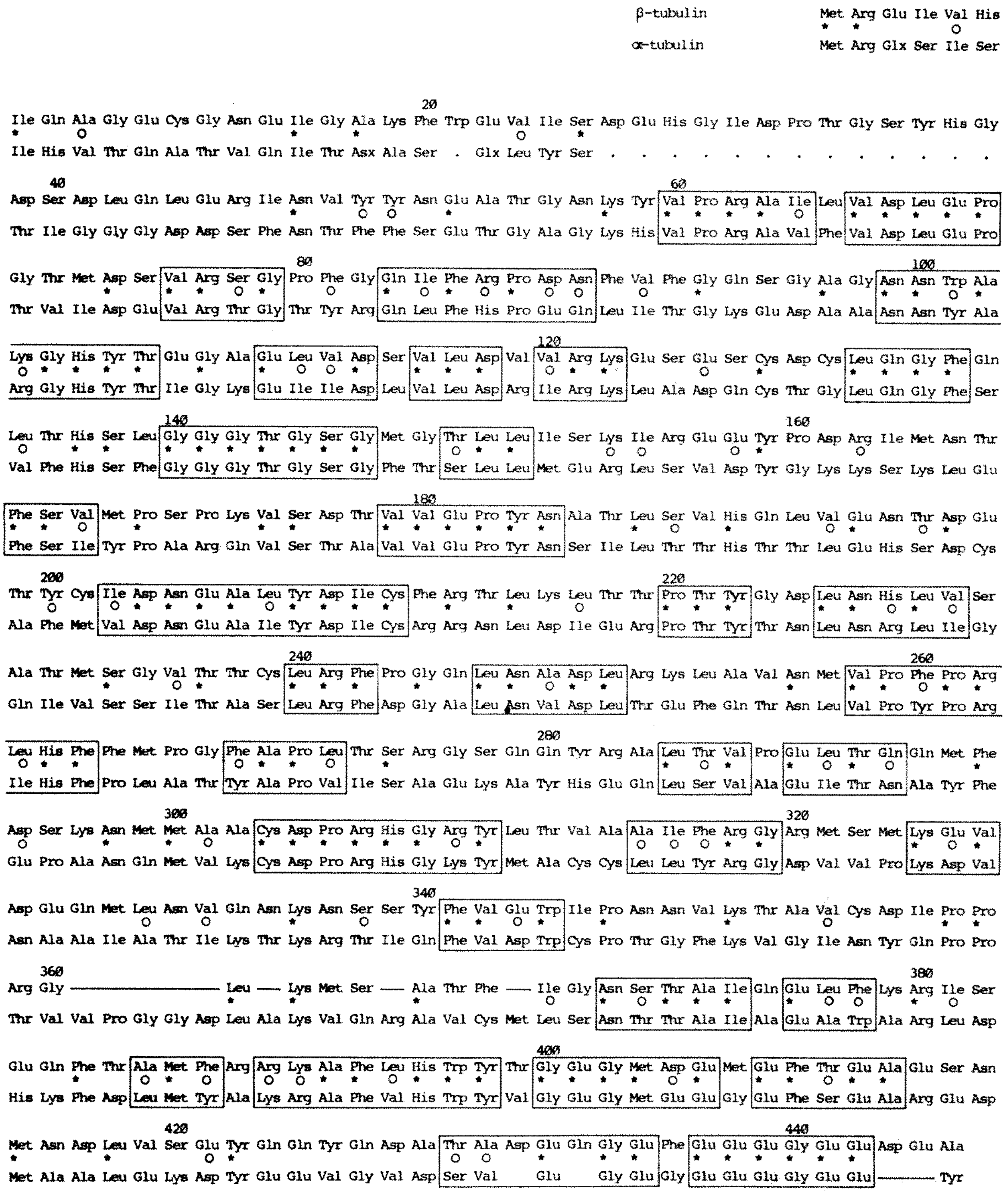

β-tubulin Met Arg Glu Ile Val His
α-tubulin Met Arg Glx Ser Ile Ser
20
Ile Gln Ala Gly Glu Cys Gly Asn Glu Ile Gly Ala Lys Phe Trp Glu Val Ile Ser Asp Glu His Gly Ile Asp Pro Thr Gly Ser Tyr His Gly
Ile His Val Thr Gln Ala Thr Val Gln Ile Thr Asx Ala Ser . Glx Leu Tyr Ser . . . .
40 60
Asp Ser Asp Leu Gln Leu Glu Arg Ile Asn Val Tyr Tyr Asn Glu Ala Thr Gly Asn Lys Tyr Val Pro Arg Ala Ile Leu Val Asp Leu Glu Pro
Thr Ile Gly Gly Gly Asp Asp Ser Phe Asn Thr Phe Phe Ser Glu Thr Gly Ala Gly Lys His Val Pro Arg Ala Val Phe Val Asp Leu Glu Pro
80 100
Gly Thr Met Asp Ser Val Arg Ser Gly Pro Phe Gly Gln Ile Phe Arg Pro Asp Asn Phe Val Phe Gly Gln Ser Gly Ala Gly Asn Asn Trp Ala
Thr Val Ile Asp Glu Val Arg Thr Gly Thr Tyr Arg Gln Leu Phe His Pro Glu Gln Leu Ile Thr Gly Lys Glu Asp Ala Ala Asn Asn Tyr Ala
120
Lys Gly His Tyr Thr Glu Gly Ala Glu Leu Val Asp Ser Val Leu Asp Val Val Arg Lys Glu Ser Glu Ser Cys Asp Cys Leu Gln Gly Phe Gln
Arg Gly His Tyr Thr Ile Gly Lys Glu Ile Ile Asp Leu Val Leu Asp Arg Ile Arg Lys Leu Ala Asp Gln Cys Thr Gly Leu Gln Gly Phe Ser
140 160
Leu Thr His Ser Leu Gly Gly Gly Thr Gly Ser Gly Met Gly Thr Leu Leu Ile Ser Lys Ile Arg Glu Glu Tyr Pro Asp Arg Ile Met Asn Thr
Val Phe His Ser Phe Gly Gly Gly Thr Gly Ser Gly Phe Thr Ser Leu Leu Met Glu Arg Leu Ser Val Asp Tyr Gly Lys Lys Ser Lys Leu Glu
180
Phe Ser Val Met Pro Ser Pro Lys Val Ser Asp Thr Val Val Glu Pro Tyr Asn Ala Thr Leu Ser Val His Gln Leu Val Glu Asn Thr Asp Glu
Phe Ser Ile Tyr Pro Ala Arg Gln Val Ser Thr Ala Val Val Glu Pro Tyr Asn Ser Ile Leu Thr Thr His Thr Thr Leu Glu His Ser Asp Cys
200 220
Thr Tyr Cys Ile Asp Asn Glu Ala Leu Tyr Asp Ile Cys Phe Arg Thr Leu Lys Leu Thr Thr Pro Thr Tyr Gly Asp Leu Asn His Leu Val Ser
Ala Phe Met Val Asp Asn Glu Ala Ile Tyr Asp Ile Cys Arg Arg Asn Leu Asp Ile Glu Arg Pro Thr Tyr Thr Asn Leu Asn Arg Leu Ile Gly
240 260
Ala Thr Met Ser Gly Val Thr Thr Cys Leu Arg Phe Pro Gly Gln Leu Asn Ala Asp Leu Arg Lys Leu Ala Val Asn Met Val Pro Phe Pro Arg
Gln Ile Val Ser Ser Ile Thr Ala Ser Leu Arg Phe Asp Gly Ala Leu Asn Val Asp Leu Thr Glu Phe Gln Thr Asn Leu Val Pro Tyr Pro Arg
280
Leu His Phe Phe Met Pro Gly Phe Ala Pro Leu Thr Ser Arg Gly Ser Gln Gln Tyr Arg Ala Leu Thr Val Pro Glu Leu Thr Gln Gln Met Phe
Ile His Phe Pro Leu Ala Thr Tyr Ala Pro Val Ile Ser Ala Glu Lys Ala Tyr His Glu Gln Leu Ser Val Ala Glu Ile Thr Asn Ala Tyr Phe
300 320
Asp Ser Lys Asn Met Met Ala Ala Cys Asp Pro Arg His Gly Arg Tyr Leu Thr Val Ala Ala Ile Phe Arg Gly Arg Met Ser Met Lys Glu Val
Glu Pro Ala Asn Gln Met Val Lys Cys Asp Pro Arg His Gly Lys Tyr Met Ala Cys Cys Leu Leu Tyr Arg Gly Asp Val Val Pro Lys Asp Val
340
Asp Glu Gln Met Leu Asn Val Gln Asn Lys Asn Ser Ser Tyr Phe Val Glu Trp Ile Pro Asn Asn Val Lys Thr Ala Val Cys Asp Ile Pro Pro
Asn Ala Ala Ile Ala Thr Ile Lys Thr Lys Arg Thr Ile Gln Phe Val Asp Trp Cys Pro Thr Gly Phe Lys Val Gly Ile Asn Tyr Gln Pro Pro
360 380
Arg Gly ——— Leu — Lys Met Ser — Ala Thr Phe — Ile Gly Asn Ser Thr Ala Ile Gln Glu Leu Phe Lys Arg Ile Ser
Thr Val Val Pro Gly Gly Asp Leu Ala Lys Val Gln Arg Ala Val Cys Met Leu Ser Asn Thr Thr Ala Ile Ala Glu Ala Trp Ala Arg Leu Asp
400
Glu Gln Phe Thr Ala Met Phe Arg Arg Lys Ala Phe Leu His Trp Tyr Thr Gly Glu Gly Met Asp Glu Met Glu Phe Thr Glu Ala Glu Ser Asn
His Lys Phe Asp Leu Met Tyr Ala Lys Arg Ala Phe Val His Trp Tyr Val Gly Glu Gly Met Glu Glu Gly Glu Phe Ser Glu Ala Arg Glu Asp
420 440
Met Asn Asp Leu Val Ser Glu Tyr Gln Gln Tyr Gln Asp Ala Thr Ala Asp Glu Gln Gly Glu Phe Glu Glu Glu Gly Glu Glu Asp Glu Ala
Met Ala Ala Leu Glu Lys Asp Tyr Glu Glu Val Gly Val Asp Ser Val Glu Gly Glu Gly Glu Glu Glu Gly Glu Glu ——— Tyr

Fig. 5 Comparison of the amino acid sequences predicted from the *α* and *β* tubulin mRNAs. The predicted protein sequences for *α* and *β* tubulin have been aligned to show regions of homology. Open circles indicate homologous, conservative substitutions. Boxed regions indicate areas of high homology. For completeness, the sequence of the first 25 amino acids of chick brain *α* tubulin deduced by Luduena and Woodward[5] is also shown. *Identical amino acids in the two sequences.

event is not the addition but the removal of tyrosine from the C-terminus. This in turn strengthens the argument that there is probably a cyclic process of detyrosylation and tyrosylation and that this may be an important factor in modulating the activity of tubulin in the cell.

Codon usage in α and β tubulin

Figure 4*a*, *b* illustrates the prevalence of codons utilized for each of the amino acids in the translated frames of α and β tubulin, respectively. Surprisingly, strong biases are observed for both tubulins in the case of codons in which the third position is occupied by a G or C. For example, the threonine codons ACC and ACG are used 28 times in both α and β sequences, whereas the ACU and ACA codons are used only twice in either sequence. This contrasts with the reported codon usage for other chicken mRNAs sequenced to date. In ovalbumin mRNA[16], the ACU and ACA codons are used 11 times, but ACC is used only 4 times and ACG not at all. Similarly, for the *src* and *env* proteins of Rous sarcoma virus[17], the known threonine codon usages are 15:18 and 3:6 for ACC+ACG:ACA+ACU. Overall, codons which have G or C in the third position are utilized in the two tubulin sequences 9.6 times more frequently for α tubulin and 8.7 times for β tubulin than codons ending in A or U. The corresponding ratios for ovalbumin, *src* and *env* are 0.9, 2.5 and 1.0. The significance of the disparate codon usage is yet unclear; however, it may reflect some constraint on tubulin mRNA sequences beyond simple conservation of protein sequence.

Homologies between α and β tubulin

The amino acid sequences which we have deduced for α and β tubulin are shown in Fig. 5. The two sequences have been aligned to highlight regions of homology. The tubulins show several interesting features. First, the two sequences are 40–55% homologous throughout, depending on whether exact homology or conservative substitutions are considered. Such homology was first suggested by Luduena and Woodward[5] when they sequenced the N-termini of each chain and found that 11 out of 25 residues were identical. This homology, though distributed among all regions of the sequences, appears in more than 15 clusters of conserved amino acids. For example, the C-terminal regions of both chains contain highly acidic runs of homologous amino acids. Residues 387–411 (numbers are with respect to the β sequence) represent a cluster of 25 amino acids of which all but three are homologous and residues 429–443 comprise a cluster of 14 amino acids of which 11 are homologous. At the nucleic acid level we also find considerable sequence homology (~50%), again with some strongly conserved clusters of nucleic acid residues. An example of this DNA homology can be found in the nucleic acid sequences coding for the amino acids 58–70. Out of 13 amino acids, 11 are homologous in this region; 33 out of 39 nucleic acid residues are identical.

Overall, our sequencing data unambiguously demonstrate that substantial sequence homology exists at both the protein and nucleic acid sequence levels for α and β tubulin. These data confirm the widely held contention that the two tubulin subunits arose from a common ancestral sequence.

As a complementary approach to determining homology we have predicted secondary structure from the amino acid sequences by the methods of Garnier *et al.*[18] and Chou and Fasman[19] for α and β tubulin. With the exception of the C-terminal regions which are both predicted to be highly α helical, additional levels of structural homology were not obvious.

Further comments

The sequence data presented here demonstrate that α and β tubulin share substantial sequence homology and thus probably derived from a common ancestor. There has been speculation that sequence homology might exist between tubulin and muscle proteins. Krauhs *et al.*[20] have reported the location of such homology between actin and α tubulin at residues 22–31 of actin and some residues of α tubulin. This portion of the actin sequence, conserved in all actins sequenced so far[21–26], is ^{22}Ala-Gly-Asp-Asp-Ala-Pro-Arg-Ala-Val-Phe; the corresponding tubulin sequence which we have deduced is ^{57}Ala-Gly-Lys-His-Val-Pro-Arg-Ala-Val-Phe. Thus 7 out of 10 residues in the two proteins, including one stretch of five, are identical. Whether this reflects true homology due to a common binding site for nucleic acid cofactors or a common accessory protein, or whether the homology is merely coincidental remains to be determined by further evolutionary studies.

The hexanucleotide sequence, AAUAAA, has been found ~20 nucleotides from the poly(A) in almost all known polyadenylated RNA sequences and is thought to be the signal for such polyadenylation. However, we find the heptanucleotide, ACAUAAA, located 22 nucleotides 5′ to the beginning of the poly(A) in our β tubulin cDNA. In addition, other mRNAs, like those encoding rat pancreatic amylase[27] and anglerfish pancreatic somatostatin[28], have recently been shown to carry the variant, AAUUAAA. These variations probably reflect flexibility in the sequence requirement of this site.

Heterogeneity of both α and β tubulin chains has been reported many times. This heterogeneity has been observed repeatedly by one- or two-dimensional gel analyses and peptide mapping, and recently by genetic analysis[29,30]. Indeed, we have demonstrated that multiple gene sequences exist for α tubulin and for β tubulin in several organisms (that is, at least four α and four β genes in chicken and *Drosophila*—see refs 3, 4). The availability of the entire protein sequence for one α tubulin and one β tubulin from embryonic chick brain should facilitate additional efforts (1) to localize the regions of heterogeneity of different tubulin subunits and (2) to explore the functional significance of each heterogeneity.

We thank Sonia Bock, David Agard and Dr H. Martinez for their assistance with the computer programs used in the DNA and protein sequence analysis. This work was supported by NIH grants GM 25157 to P.V. and GM 26875 to M.K.

Received 20 October; accepted 2 December 1980.

1. Roberts, K. & Hyams, J. S. (eds) *Microtubules* (Academic, New York, 1979).
2. Kirschner, M. W. *Int. Rev. Cytol.* **54,** 1–71 (1979).
3. Cleveland, D. W. *et al. Cell* **20,** 95–105 (1980).
4. Sanchez, F., Natzle, J., Cleveland, D. W., Kirschner, M. W. & McCarthy, B. J. *Cell* **22,** 845–854 (1980).
5. Luduena, R. F. & Woodward, D. O. *Proc. natn. Acad. Sci. U.S.A.* **70,** 3594–3598 (1973).
6. Lu, R. C. & Elzinga, M. *Biochim. biophys. Acta* **537,** 320–328 (1978).
7. Ponstingl, H., Little, M., Krauhs, E. & Kempf, T. *Nature* **282,** 423–424 (1979).
8. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
9. Feit, H., Slusarek, L. & Shelanski, M. L. *Proc. natn. Acad. Sci. U.S.A.* **68,** 2028–2032 (1971).
10. Cleveland, D. W., Kirschner, M. W. & Cowan, N. J. *Cell* **15,** 1021–1031 (1978).
11. Arce, C. A., Rodriguez, J. A., Barra, H. S. & Caputto, R. *Eur. J. Biochem.* **59,** 145–149 (1975).
12. Argarana, C. E., Arce, C., Barra, H. S. & Caputto, R. *Archs Biochem. Biophys.* **180,** 264–268 (1977).
13. Raybin, D. & Flavin, M. *Biochem. biophys. Res. Commun.* **65,** 1088–1095 (1975).
14. Raybin, D. & Flavin, M. *Biochemistry* **16,** 2189–2194 (1977).
15. Rodriguez, J. A. & Borisy, G. G. *J. Cell Biol.* **75,** 296a (1978).
16. McReynolds, L. *et al. Nature* **273,** 723–728 (1978).
17. Czernilofsky, A. P., Levinson, A. D., Varmus, H. E. & Bishop, J. M. *Nature* **287,** 198–203 (1980).
18. Garnier, J., Osguthorpe, D. J. & Robson, B. *Molec. Biol.* **120,** 97–120 (1978).
19. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13,** 222–245 (1974).
20. Krauhs, E., Little, M., Kempf, T. & Ponstingl, H. *Eur. J. Biochem.* **22,** 285 (1980).
21. Collins, J. H. & Elzinga, M. *J. biol. Chem.* **250,** 5906–5920 (1975).
22. Lu, R. C. & Elzinga, M. *Biochemistry* **16,** 5801–5806 (1977).
23. Vandekerchove, J. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1106–1110 (1978).
24. Vandekerchove, J. & Weber, K. *Nature* **276,** 720–721 (1978).
25. Gallwitz, D. & Sures, I. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2546–2550 (1980).
26. Ng, R. & Abelson, J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3912–3916 (1980).
27. MacDonald, R. J. *et al. Nature* **287,** 117–122 (1980).
28. Hobart, P., Crawford, R., Shen, L.-P., Pictet, R. & Rutter, W. J. *Nature* **288,** 137–141 (1980).
29. Kemphues, K. J., Raff, R. A., Kaufman, T. C. & Raff, E. C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3991–3995 (1979).
30. Kemphues, K. J., Raff, E. C., Raff, R. A. & Kaufman, T. C. *Cell* **21,** 445–454 (1980).

LETTERS

Can negative molecular ions be detected in dense interstellar clouds?

Eric Herbst

Department of Physics, Duke University, Durham,
North Carolina 27706

The recent laboratory measurement[1-3] of rapid radiative electron attachment processes

$$\mathrm{A + e \rightarrow A^{-}}h\nu$$

where A is a molecular species has renewed speculation on whether negative molecular ions can be synthesized efficiently in dense interstellar clouds. We argue here that for certain interstellar species A, the abundance ratio [A⁻]/[A] may be as high as 0.01–0.1 in commonly assumed physical conditions. If this abundance ratio were correct, negative molecular ions might be detectable in dense interstellar clouds if their microwave spectral frequencies had been determined in the laboratory. It will be shown, however, that this is currently an unlikely prospect.

In recent studies[4-6] of the gas phase chemistry of 'dense' interstellar clouds, the major processes considered in the ambient conditions of low density (10^4–10^6 cm^{-3}) and low temperature (10–50 K) are reactions between positive ions and molecules. These produce polyatomic ions and, on reaction with electrons, a variety of neutral molecules. Cosmic-ray penetration of the clouds produces both an initial assortment of positive ions and electrons. Although the subsequent formation of negative ions through radiative attachment of electrons has been considered[7], it has been concluded that negative ions are both scarce and only a minor contributor to ion–molecule synthetic processes. The scarcity results from the generally slow rate of radiative attachment.

Consider a molecular species A: in dense interstellar clouds, the negative ion $\mathrm{A^-}$ can be formed and depleted by the reaction sequence

$$\mathrm{A + e} \xrightarrow{k_{\mathrm{ratt}}} \mathrm{A^{-}} + h\nu \qquad (1)$$

$$\mathrm{A^{-} + B} \xrightarrow{k_2} \mathrm{AB + e,\ C^{-} + D} \qquad (2)$$

$$\mathrm{A^{-}} + h\nu \xrightarrow{k_3} \mathrm{A + e} \qquad (3)$$

$$\mathrm{A^{-} + I^{+}} \xrightarrow{k_4} \text{Products} \qquad (4)$$

where B and $\mathrm{I^+}$ represent generic neutral and positively-charged species, respectively. If $\mathrm{A^-}$ is at steady-state, the abundance ratio $[\mathrm{A^-}]/[\mathrm{A}]$ is

$$\frac{[\mathrm{A^{-}}]}{[\mathrm{A}]} = \frac{k_{\mathrm{ratt}} f_{\mathrm{e}}}{k_2 f_{\mathrm{B}} + k_3/n + k_4 f_{\mathrm{I^{+}}}} \qquad (5)$$

where n is the total gas density and f_{i} refers to the fractional abundance of species i. Typical values of k_2 are $\sim 10^{-9}$ cm^{-3} s^{-1} (ref. 8) (independent of temperature) and, of k_4, are $\sim 10^{-7}$–10^{-8} cm^3 s^{-1} at room temperature with $T^{-1/2}$ dependence[9]. The rate coefficient k_3 for photodetachment can generally be neglected in dense clouds for negative ions of several electron volts stability. High pressure, saturated three-body attachment rate coefficients have been studied extensively for stable neutral species[10] and found to vary widely—these values are useful upper limits to the radiative attachment coefficients. For most neutral species, high pressure attachment coefficients are considerably below collision rate coefficients ($\sim 10^{-7}$ cm^3 s^{-1}) at room temperature. High pressure attachment and measured autodetachment rates[10] can lead to an estimate for radiative attachment rate coefficients; such estimates show that for most neutral species these coefficients are quite small at room temperature. Therefore, the recent measurement of 'rapid' (~0.1–0.01 of the collision rate) room temperature radiative attachment coefficients at low pressure by the ICR technique is indeed intriguing[1,2]. There are no laboratory attachment and detachment data for most interstellar neutral species capable of forming negative ions, thus a theoretical treatment is required. A simple theory for k_{ratt} can be derived from phase space considerations[11-13]. Assuming that no vibrational excitation of the neutral species need be considered, the theory reduces to the following expression for k_{ratt}:

$$k_{\mathrm{ratt}}\,(\mathrm{cm^3\,s^{-1}}) \sim \frac{10^3 h^3 (kT)^{-1/2}}{(2\pi m_{\mathrm{e}})^{3/2}} \times \frac{g_{\mathrm{A^-}}}{2g_{\mathrm{A}}} N_{\mathrm{v}}(\mathrm{EA}) \qquad (6)$$

where the computed value of k_{ratt} cannot exceed the (unknown) rate coefficient for temporary electron capture. In equation (6) h is Planck's constant; m_{e}, the electron mass; $N_{\mathrm{v}}(\mathrm{EA})$, the vibrational density of states of $\mathrm{A^-}$ at an internal energy equal to the electron affinity; k, the Boltzmann constant; and g_{A} and $g_{\mathrm{A^-}}$, the electronic degeneracy factors of A and $\mathrm{A^-}$. Equation (6) is capable of reproducing the large experimental value[1,2] of $k_{\mathrm{ratt}} = 2\times10^{-8}$ cm^3 s^{-1} at 300 K for $\mathrm{SF_6}$ attachment if, as has been assumed[11], the vibrational frequencies of $\mathrm{SF_6^-}$ are half those of $\mathrm{SF_6}$. In general, the lack of known negative-ion vibrational frequencies limits the usefulness of equation (6) even if the statistical assumptions embodied in it are correct.

According to equation (6), k_{ratt} will be large when $N_{\mathrm{v}}(\mathrm{EA})$ is large. The vibrational density of states is a severe function of internal energy and molecular size[12,13] and increases greatly with both increasing electron affinity and molecular size. Based on equation (6) and a knowledge of interstellar chemistry, the most abundant interstellar molecular negative ions stem from neutral free radicals with estimated large (>2eV) electron affinities. For a variety of these species, such as $\mathrm{C_4H}$, $\mathrm{C_5N}$, $\mathrm{C_7N}$, $\mathrm{C_9N}$, $\mathrm{C_2H_2CN}$, $\mathrm{C_2H_3O}$, $\mathrm{C_2H_3O_2}$ and possibly $\mathrm{C_3N}$, equation (6) can be used to make plausible radiative attachment rate coefficients near the collision limit of 10^{-7} cm^3 s^{-1} at cloud temperatures of 10–50 K. The first and last radicals mentioned have been detected in clouds[14,15] while the others are possible on synthetic grounds[4-6]. The list has been prepared by removing hydrogen atoms from known stable interstellar molecules. Negative ions formed from such radicals are likely to react with atomic hydrogen via associative detachment processes, for example:

$$\mathrm{C_3N^{-} + H \longrightarrow IHC_3N + e}$$

Assuming that the fractional abundance of atomic hydrogen in dense clouds is $f_{\mathrm{B}} \sim 10^{-4}$ (refs 4–6), k_{ratt} is $\leqslant 10^{-7}$ cm^3 s^{-1}, k_4 is $\sim 10^{-6}$ cm^3 s^{-1}, k_2 is $\sim 10^{-9}$ cm^3 s^{1}, and $f_{\mathrm{I^+}} \simeq f_{\mathrm{e}}$, we obtain from equation (5) that

$$[\mathrm{A^{-}}]/[\mathrm{A}] \leqslant 10^6 f_{\mathrm{e}}/(1 + 10^7 f_{\mathrm{e}}) \qquad (7)$$

Estimates for f_{e} in dense clouds range from 10^{-7}–10^{-8} (refs 4–6, 16). For $f_{\mathrm{e}} \geqslant 10^{-7}$, equation (7) predicts $[\mathrm{A^-}]/[\mathrm{A}]$ to be $\leqslant 0.1$. For the more conservative value of $f_{\mathrm{e}} \sim 10^{-8}$, $[\mathrm{A^-}]/[\mathrm{A}] \leqslant 0.01$. Note that for equation (7) to apply, $[\mathrm{A^-}] < [\mathrm{e}]$.

The major uncertainty in our argument is whether or not radiative attachment rate coefficients near the collision-limited value of $\sim 10^{-7}$ cm^3 s^{-1} occur for only a few measured species[1,2] mostly larger than known interstellar molecules, or whether some smaller radicals with large electron affinities also possess such large values of k_{ratt}. The above simple theory supports the latter possibility. Still, negative molecular ions based on such radicals are not expected to be overly abundant. If we take as

0028-0836/81/080656—02$01.00

typical an abundance ratio[4-6] [A]/[AH] of $\leqslant 10^{-2}$ where AH is the stable species (for example, $A = C_5N$, $A^- = C_5N^-$, $AH = HC_5N$), then $[A^-]/[AH] \leqslant 10^{-3} - 10^{-4}$. Experimental work on the attachment of small free radicals is desirable to determine whether 10^{-3}–10^{-4} is a reasonable estimate or a severe upper limit. If the former assessment is correct, certain negative ions may eventually be detectable in dense interstellar clouds. Also if some relevant radiative attachment rates do approach 10^{-7} cm^3 s^{-1}, these and related processes must be included in detailed cloud models.

This research was supported by NSF grants AST-7904529 and AST-8020321.

Received 9 October; accepted 12 December 1980.

1. Woodin, R. L., Foster, M. S. & Beauchamp, J. L. *J. chem. Phys.* **72,** 4223–4227 (1980).
2. Foster, M. S. & Beauchamp, J. L. *Chem. phys. Lett.* **31,** 482–485 (1975).
3. Christodoulides, A. A. & Christophorou, L. G. *Chem. phys. Lett.* **61,** 553–557 (1979).
4. Herbst, E. & Klemperer, W. *Astrophys. J.* **185,** 505–533 (1973).
5. Mitchell. G. F., Ginsburg, J. L. & Kuntz, P. J. *Astrophys. J. Suppl.* **38,** 39–68 (1978).
6. Prasad, S. S. & Huntress, W. T. Jr *Astrophys. J. Suppl.* **43,** 1–36 (1980).
7. Dalgarno, A. & McCray, R. A. *Astrophys. J.* **181,** 95–100 (1973).
8. Albritton, D. L. *Atomic Data and Nuclear Tables* **22,** 1–101 (1978).
9. Smith, D. & Adams, N. G. *Top. curr. Chem.* **89,** 1–43 (1980).
10. Christophorou, L. G. *Adv. Electron. Electron Phys.* **46,** 55-129 (1978).
11. Klots, C. E. *Chem. phys. Lett.* **38,** 61–63 (1976).
12. Bass, L., Chesnavich, W. J. & Bowers, M. T. *J. Am. chem. Soc.* **101,** 5493–5502 (1979).
13. Herbst, E. *Astrophys. J.* **237,** 462–470 (1980).
14. Guélin, M., Green, S. & Thaddeus, P. *Astrophys. J. Lett.* **224,** L27–L30 (1978).
15. Guélin, M. & Thaddeus, P. *Astrophys. J. Lett.* **212,** L81–86 (1977).
16. Langer, W. D., Wilson, R. W., Henry, P. S. & Guélin, M. *Astrophys. J. Lett.* **225,** L139–L142 (1978).

Massive neutrinos, helium production and the primordial magnetic field

Stuart L. Shapiro & Ira Wasserman

Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853

Experimental evidence[1] for a non-zero neutrino rest mass, $m_\nu \neq 0$, while still controversial[2], poses intriguing cosmological questions[3-10]. One reassurance is that the possibility of non-zero masses does not necessarily alter previous predictions of the standard big-bang model for the primordial ^{4}He abundance, Y. We show here that a non-zero neutrino rest mass, $m_\nu \sim 10$ eV, may provide the most stringent upper limit proposed to date on the strength of any ordered, primordial magnetic field: $B_0 \leqslant 4\times10^{-9}/m_\nu$ (eV) G. This limit is set by requiring that previous predictions of the standard big bang model for the primordial ^{4}He abundance remain unaltered by the spin procession and subsequent thermalization of right-handed, as well as left-handed, Dirac neutrinos in the early Universe.

Massive neutrinos can be either Majorana or Dirac fermions. Majorana neutrinos are represented by two-component spinor wave functions that are CP eigenstates in the neutrino rest frame. As Majorana mass terms do not conserve lepton number there is no Majorana antineutrino. The total number of spin degrees of freedom per Majorana neutrino flavour is, therefore, $g = 2$, thus the same as for massless neutrinos. Dirac neutrinos, like electrons, are represented by four-component spinor wave functions that are C and P eigenstates in the neutrino rest frame. Dirac mass terms conserve lepton number, so there are both Dirac neutrinos and antineutrinos, and $g = 4$.

For Dirac neutrinos with $m_\nu \neq 0$, both left- and right-handed neutrinos could be present in the early Universe, thereby doubling the number of possible neutrino types. However, neutrino–fermion weak reactions mediated either by vector ($\mathbf{W}^\pm$, $\mathbf{Z}^0$) or scalar (Higgs) bosons through the standard Weinberg–Salam–Glashow[11-13] (WSG) model, extended to include Dirac neutrinos with $m_\nu \neq 0$, cannot maintain ν_R in thermal equilibrium at the critical decoupling epoch, $kT_D \sim 1$ MeV (refs 9, 10). In fact, ν_R decouple much earlier at $kT \gg 80$ GeV. Previous studies have shown[14] that, for a present-day baryon density $\rho_B \gtrsim 10^{-2}\rho_c$, any new relativistic particles in the early Universe must decouple at $kT \gtrsim 0.2$ GeV to satisfy the observational constraint on the ^{4}He abundance[15], $Y \leqslant 0.25$. So, as far as neutrino–fermion weak interactions are concerned, the WSG model extended to include massive neutrinos is consistent with this requirement.

A further consequence of non-zero neutrino rest masses is that neutrinos acquire a magnetic moment ($G_F \equiv 10^{-5} m_p^{-2}$),

$$\mu_\nu = \frac{3eG_F\hbar m_\nu}{8\sqrt{2}\pi^2 c} \tag{1}$$

from diagrams of electro-weak order[16,17]. Consequently, neutrinos can precess in an external magnetic field. Extremely relativistic neutrinos in the early Universe can flip helicities at a rate[16,17]

$$\Gamma_{LR}^{B} \approx \frac{2}{\pi}\frac{\mu_\nu B_\perp}{\hbar} \approx \frac{3eG_F m_\nu B_\perp}{4\sqrt{2}\pi^3 c} \tag{2}$$

where $B_\perp$ is the component of the primordial field perpendicular to the neutrino velocity. For a sufficiently large magnetic field, $\Gamma_{LR}^{B} \gtrsim \Gamma_{LL}$, where Γ_{LL} is the rate of neutrino–fermion interactions which thermalize the left-handed neutrinos. For such rapid flipping an equilibrium population of right-handed neutrinos can result, doubling the number of thermal (Dirac) neutrino types (from $N_\nu \gtrsim 3$ to $2N_\nu \gtrsim 6$). This would violate the constraint[14,15] of $N_\nu \leqslant 4$ imposed on the standard big-bang model by the observed ^{4}He abundance, $Y \leqslant 0.25$, for $\rho_B \gtrsim 10^{-2}\rho_c$. (If $\rho_B \ll 10^{-2}\rho_c$ the allowed number of neutrino types $N_\nu \gg 4$; a conservative lower limit[5] is $\rho_B \gtrsim 0.006\rho_c$.) Alternatively, satisfying the constraint $N_\nu \leqslant 4$ by requiring

$$\Gamma_{LR}^{B} \leqslant \Gamma_{LL} \tag{3}$$

up to ν_L-decoupling at $T_D \sim 1$ MeV places a surprisingly stringent limit on the strength B_0 of any present-day relic magnetic field.

Our limit is an immediate consequence of three basic assumptions:

(1) The standard big-bang model of nucleosynthesis in the early Universe[18,19] is valid and the primordial ^{4}He abundance satisfies[15] $Y \leqslant 0.25$.

This assumption requires inequality (3) to hold throughout the nucleosynthesis epoch. Enforcing this requirement at T_D, when $\Gamma_{LL} \sim \Gamma_{exp}$, where

$$\Gamma_{exp} \sim (G\rho)^{1/2} \sim \frac{(kT)^2}{\hbar^{3/2}c^{5/2}}$$

is the expansion rate and ρ is the total density of the Universe, gives

$$m_\nu B_D \leqslant \frac{4\pi^3\sqrt{2}}{3}\frac{G^{1/2}(kT_D)^2}{eG_F c^{3/2}\hbar^{3/2}}. \tag{4}$$

In equation (4) we have set $B_\perp \sim B_D$, the field strength at neutrino decoupling.

(2) The magnetic field remains 'frozen-in' to the perfectly conducting matter so that $BR^2 \sim$ constant, where R is the expansion parameter. Hence, BT^{-2} does not change as the Universe expands.

Because $\Gamma_{LR}^{B} \propto B \propto T^2$ while[18,19] $\Gamma_{LL} \propto T^5$, inequality (4) ensures that the population of ν_R cannot equilibrate for any $T > T_D$. Furthermore, if T_0 is the present temperature of the microwave background, then $B_0 = B_D(T_0/T_D)^2$, so inequality (4) gives

$$m_\nu B_0 \leqslant \frac{4\pi^3\sqrt{2}}{3}\frac{G^{1/2}(kT_0)^2}{eG_F c^{3/2}\hbar^{3/2}} \approx 3.9\times10^{-9} \quad \text{eVG} \tag{5}$$

for $T_0 = 2.7$ K (refs 18, 19).

0028-0836/81/080657—02$01.00

(3) The neutrino mass probably lies in the range

$$4 \leqslant m_\nu(\text{eV}) \leqslant 20 \qquad (6)$$

It has been argued[3-5] that neutrinos with masses in this range would solve the missing mass problem for clusters of galaxies without demanding unacceptably high mass-to-light ratios for individual and binary galaxies and small groups. For this mass range, inequality (5) implies

$$B_0 \leqslant \frac{4\pi^3\sqrt{2}}{3} \frac{G^{1/2}(kT_0)^2}{eG_F m_\nu c^{3/2}\hbar^{3/2}} \approx (0.2\text{–}1.0)\times 10^{-9} \quad \text{G} \qquad (7)$$

The right-hand side of inequality (7) is an interesting combination of fundamental microphysical and cosmological constants. Unlike many comparable expressions—such as the large numbers—the sole cosmological parameter, T_0, is not a major source of uncertainty in equation (7).

In deriving inequality (7) we have implicitly assumed that the magnetic field is uniform over all length scales $L \leqslant L_D \equiv ct_D \sim 3\times 10^{10}$ cm at decoupling, implying uniformity over length scales $L_0 \leqslant L_D z_D \sim 30$ pc at present. In the unlikely event that the field is actually tangled over smaller scales $l_0 \ll L_0$ today (and l_0/z_D at decoupling) the upper limit to B_0 is increased over the value given in equation (7) by the random-walk factor $(L_0/l_0)^{1/2} = (30\text{pc}/l_0)^{1/2}$.

Note that inequality (7) provides the most stringent limit on B_0 proposed to date. For example, invoking assumption (2) we find that the requirement that the magnetic energy density does not exceed the radiation energy density only implies

$$B_0 < (8\pi a)^{1/2} T_0^2 \approx 3.2 = 10^{-6} \quad \text{G} \qquad (8)$$

for a present black-body radiation temperature $T_0 \approx 2.7$ K. Because $B^2/\rho_B \propto T$, where ρ_B is the baryon mass density, inequality (10) also guarantees that the magnetic energy density remains below the baryon mass density throughout the matter-dominated era. Such a magnetic field is also much too weak to produce an observable anisotropy in the microwave background[20-22]. To guarantee $B^2/8\pi at^4 < 1$ all the way back to the very earliest epochs, assumption (2), leading to the inequality (8), must be modified by a logarithmic correction factor[23]. Thus, consider a Bianchi type I model with a uniform magnetic field for the anisotropic expansion of a perfectly conducting universe. Requiring that $B^2/8\pi aT^4 < 1$ back to the Planck time $t_p \sim (\hbar G/c^5)^{1/2} \sim 10^{43}$s demands[23]

$$B_0 < (1\text{–}3)\times 10^{-7} \quad \text{G} \qquad (9)$$

which is still far less restrictive than equation (7).

For $T \geqslant 1$ MeV, magnetic field strengths up to $\sim 10^2 \times (8\pi a)^{1/2} T^2 \sim 10^{15} T^2$ (MeV) cause at most a ~15% change in neutron producing and destroying weak reactions involving relativistic electrons and neutrinos[24]. Accordingly, ignoring spin precession, the major effect of a primordial B-field on cosmological nucleosynthesis arises from its influence on the expansion rate and anisotropy of the Universe. Previous limits on B_0 derived by assuming the validity of the standard big-bang model for cosmological nucleosynthesis with $m_\nu \equiv 0$ are therefore equivalent to inequalities (8) and (9).

Finally, if all or part of the Faraday rotation of a distant extragalactic radio source at redshift z can be attributed to a hypothetical ionized intergalactic medium with free electron density $n_{e,0}$ one gets a limit on $n_{e,0} B_0$ if the field is uniform[25-27], and on $n_{e,0}\sqrt{l_0}$ if the field is twisted with a coherence length l_0 (refs 28–30). In the case of a uniform field one obtains a rotation measure[27]

$$\text{RM} = \frac{e^3 n_{e,0} B_{\parallel,0}}{3\pi m_e^2 c^3 H_0 \Omega^2}[(1+\Omega z)^{3/2} - 3(1-\Omega)(1+\Omega z)^{1/2} - 3\Omega + 2] \qquad (10)$$

where $\Omega \equiv 8\pi G\rho_0/3H_0^2$, ρ_0 is the total mass density, $H_0 = 100h$ km s^{-1} Mpc^{-1} is Hubble's constant, and $B_{\parallel,0}$ is the line-of-sight component of the field. Observations[30] imply an intergalactic rotation measure $\leqslant 10$ rad m^{-2}, from which one gets (assuming $\Omega z \ll 1$)[25-27]

$$B_0 \leqslant 10^{-10}\Omega_e^{-1} h^{-1} \quad \text{G} \qquad (11)$$

for a uniform field out to $z \sim 2$, and

$$B_0 \leqslant 3\times 10^{-9}\Omega_e^{-1} h^{-1/2}[l_0(\text{Mpc})]^{-1/2} \quad \text{G} \qquad (12)$$

for a twisted field[30], where $\Omega_e \equiv 8\pi G n_{e,0} m_p/3H_0^2$. Nucleosynthesis calculations, assuming $N_\nu \geqslant 3$ and $Y \leqslant 0.25$, require $\Omega_B h^2 \leqslant 0.014$, where Ω_B is the present baryon density in units of the closure density[5,15]. Taking $\Omega_e \sim \Omega_B \sim 10^{-2} h^{-2}$, inequalities (11) and (12) imply

$$B_0 \leqslant 10^{-8} h \quad \text{G} \qquad (13)$$

for a uniform field, and

$$B_0 \leqslant 3\times 10^{-7} h^{1/2}[l_0(\text{Mpc})]^{-1/2} \quad \text{G} \qquad (14)$$

for a twisted field. Smaller electron densities give less stringent limits.

Comparing inequality (7) to inequalities (8), (9), (13) and (14) shows that the most severe constraint on any primordial magnetic field is a simple consequence of a finite neutrino rest mass. Specifically, the calculated magnetic moment[16,17], μ_ν, of massive, four-component Dirac neutrinos in the WSG $SU(2)_L \times U(1)$ theory of weak and electromagnetic interactions[11-13] imposes an upper limit on B_0 an order of magnitude or more lower than any previous limit. (Two-component Majorana neutrinos do not alter the standard nucleosynthesis calculations even when $B_0 \neq 0$ so they set no useful limit to B_0.) Even more stringent limits on B_0 could arise in manifestly left–right symmetric theories with spontaneous parity non-conservation[31], even if the neutrino rest masses are very small, $m_\nu \ll 1$ eV.

We thank G. Feinberg, B. W. Lynn, and K. Olive for informative correspondence, E. E. Salpeter and S. A. Teukolsky for useful discussions, and D. Stewart for aid in preparing the manuscript. This research was supported in part by NSF grant AST 78-20708 at Cornell University.

Received 15 September; accepted 22 December 1980.

1. Reines, F., Sobel, H. W. & Pasierb, E. Preprint (UCI-10 P19–144, 1980).
2. Feynman, R. P. & Vogel, P. *Caltech Note* (1980).
3. Cowsik, R. & McClelland, J. *Phys. Rev. Lett.* **29,** 669 (1972).
4. Tremaine, S. & Gunn, J. E. *Phys. Rev. Lett.* **42,** 407 (1979).
5. Schramm, D. N. & Steigman, G. *Astrophys. J.* (submitted).
6. Bond, J. R., Efstathiou, G. & Silk, J. Preprint (University of California, Berkeley, 1980).
7. Sato, H. & Takahara, F. Preprint (Kyoto University, 1980).
8. Wasserman, I. *Astrophys. J.* (submitted).
9. Dolgov, A. D. & Zel'dovich, Ya. B. Preprint (Space Institute, Moscow, 1980).
10. Shapiro, S. L., Teukolsky, S. A. & Wasserman, I. *Phys. Rev. Lett.* **45,** 669 (1980).
11. Weinberg, S. *Phys. Rev. Lett.* **19,** 1264 (1967).
12. Salam, A. *Proc. 8th Nobel Symp.* (ed. N. Svartholm) (Almquist and Wiksells, Stockholm, 1968).
13. Glashow, S. L. *Nucl. Phys.* **22,** 579 (1961).
14. Steigman, G., Olive, K. A. & Schramm, D. N. *Phys. Rev. Lett.* **43,** 239 (1979).
15. Yang, J., Schramm, D. N., Steigman, G. & Rood, R. T. *Astrophys. J.* **227,** 697 (1979).
16. Lynn, B. W. & Feinberg, G. Preprint (CU-TP-181, 1980).
17. Fujikawa, K. & Shrock, R. Preprint (ITP-SB-80-38, 1980).
18. Peebles, P. J. E. *Physical Cosmology* (Princeton University Press, 1971).
19. Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).
20. Thorne, K. S. *Astrophys. J.* **148,** 51 (1967).
21. Jacobs, K. C. *Astrophys. J.* **153,** 661 (1968).
22. Jacobs, K. C. *Astrophys. J.* **155,** 379 (1969).
23. Zel'dovich, Ya. B. *Soviet Astr.-AJ* **13,** 608 (1970).
24. Matese, J. J. & O'Connell, R. F. *Astrophys. J.* **160,** 451 (1970).
25. Sofue, Y., Fujimoto, M. & Kawabata, K. *Publ. astr. Soc. Jap.* **20,** 388 (1968).
26. Kawabata, K., Fujimoto, M., Sofue, Y. & Fukui, M. *Publ. astr. Soc. Jap.* **21,** 293 (1969).
27. Brecher, K. & Blumenthal, G. *Astrophys. Lett.* **6,** 169 (1970).
28. Rees, M. J. & Rheinhardt, M. *Astr. Astrophys.* **19,** 189 (1972).
29. Kronberg, P. P. & Simard-Normandin, M. *Nature* **263,** 653 (1976).
30. Kronberg, P. P., Rheinhardt, M. & Simard-Normandin, M. *Astr. Astrophys.* **61,** 771 (1977).
31. Bég, M. A. B., Budny, R. V., Mohapatra, R. & Sirlin, A. *Phys. Rev. Lett.* **38,** 1252 (1977).

Density evolution of radio sources fuelled by stellar mass loss

M. E. Bailey & J. MacDonald

Astronomy Centre, University of Sussex, Falmer,
Brighton BN1 9QH, UK

At early cosmological epochs the comoving density of radio sources was very much higher than it is at present. Here we outline a theory for the evolution of the comoving density of radio sources, based on the hypothesis that mass loss from stars flows inwards in massive galaxies to fuel nuclear activity. We use a sufficient condition for the occurrence of inflow to show that the critical galaxy mass above which a total outflow becomes impossible decreases towards earlier epochs. We assume that the comoving density of radio sources is proportional to the comoving density of galaxies with inflow. The predicted density evolution of the radiogalaxy phenomenon is in good agreement with observations.

The evolution of stellar mass loss in galaxies heated by supernovae has been widely discussed[1–5]. In the special case of spherical symmetry, it can be shown[5] that a necessary condition for steady outflow beyond a stagnation radius r_s is

$$\alpha_{sn}E_{sn}(M_T - M_*(r_s)) \geq 2\pi r_s^3 \alpha P_*(r_s) + \alpha M_*(r_s)\left[\frac{GM_*(r_s)}{r_s} + \phi_*(r_s)\right] + \frac{3}{2}\alpha\int_{M_*}^{M_T}\frac{GM_*}{r}\,dM_* \qquad (1)$$

where α_{sn} is the specific mass loss rate due to supernovae, E_{sn} is the energy per unit mass of supernovae ejecta, M_T is the total mass of the galaxy, α is the specific stellar mass loss rate, $P_*(r) = \rho_* V^2/3$ is the effective pressure of the stellar velocity distribution and $\phi_*(r)$ is the gravitational potential. Adopting a de Vaucouleurs $r^{1/4}$ law[6] for the projected mass distribution of elliptical galaxies equation (1) becomes, in dimensionless form,

$$C \equiv \frac{\alpha_{sn}E_{sn}}{\alpha}\,\frac{r_e}{GM_T} \geq F_{crit}(x) \qquad (2)$$

where $x = r/r_e$ and r_e, is the effective radius. The function F_{crit} is shown for the de Vaucouleurs law in Fig. 1; it decreases monotonically from the value 0.5042 at the centre, to zero at infinity.

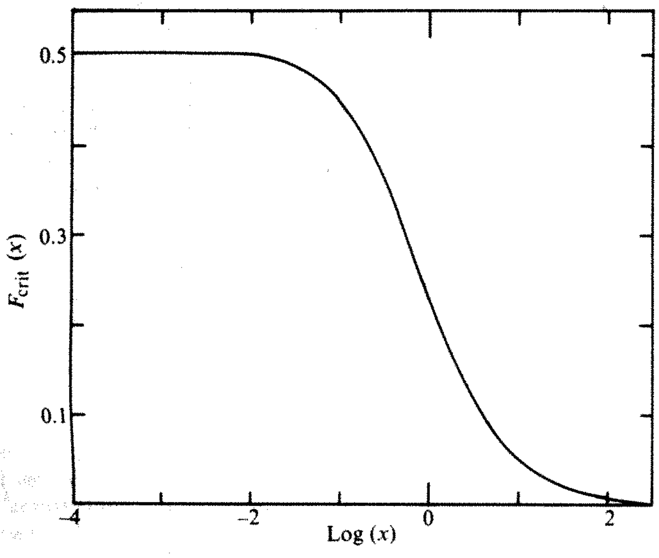


Fig. 1 The wind criterion function $F_{crit}(x)$ for the de Vaucouleurs $r^{1/4}$-law galaxy model is shown versus log (x), where $x = r/r_e$.

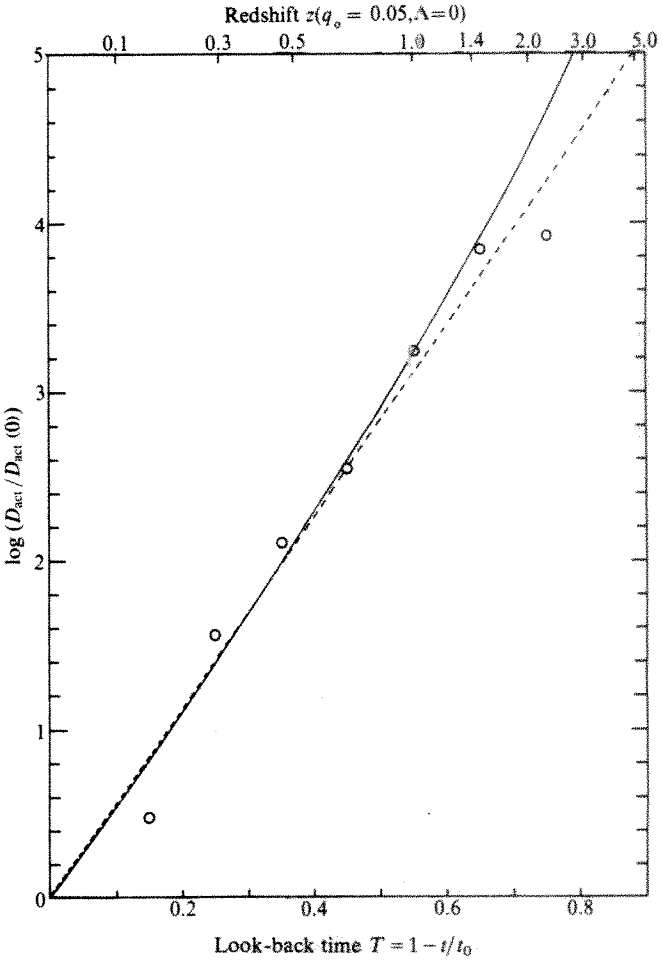


Fig. 2 The predicted comoving density of quasars is shown compared with the observations (○) and approximate analytic form (dashed line) given by Turner[19]. The predicted curve assumes $C_0 = 1.5$, $H_0 = 85$ km s^{-1} Mpc^{-1} and the luminosity function of Kirshner *et al.*[18].

We assume here that inflow occurs whenever conditions are such that total outflow (as indicated by equation (2) with $x = 0$) is not possible. This assumption favours outflowing winds and thus slightly underestimates the proportion of galaxies with inflow.

For practical use it is convenient to rewrite the quantity C on the left-hand side of equation (2) in terms of observable galaxy parameters. The specific mass loss rate $\alpha(t)$ can be obtained from a stellar population model of the galaxy[7]. For a Salpeter initial mass function we find[5] that a good approximation to $\alpha(t)$ is given by $\alpha \simeq 10^{-19}$ $(10^{10}\,\mathrm{yr}/t)\,\mathrm{s}^{-1}$. In principle, the specific supernova mass loss rate α_{sn} could also be obtained from the same stellar population model. In practice, however, uncertainties[8] in the nature of Type I supernovae make such a calculation of dubious validity. Here Type I supernovae are assumed to have low-mass progenitors[9] and are caused by mass transfer in low-mass binary systems[10]. In this case the supernova rate will change on the evolutionary time scale associated with such systems ($\sim 10^{10}$ yr). The difficulties in providing realistic estimates for the evolution of α_{sn} lead us to assume that α_{sn} is independent of epoch, consistent with the expectation of only slow change. Similarly we assume that the mean energy which Type I supernovae inject into the interstellar medium is also independent of epoch. Using the observed supernova rate for elliptical galaxies[11] $\nu_{sn} = 0.157\,h^2$ supernovae per $10^{10}\,L_\odot$ per 100 yr (where $h = H_0/50$ km s^{-1} Mpc^{-1}), and mean supernova energy 4×10^{43} J (ref. 12), we have $\alpha_{sn}E_{sn} = 10^{-8}(L_T/10^{10}L_\odot)(10^{11}M_\odot/M_T)h^2$ J kg^{-1} s^{-1}.

We now express M_T and r_e/GM_T in terms of galaxy luminosity. An empirical relation connecting the B surface

0028-0836/81/080659—03$01.00

brightness of a galaxy at r_e with the effective radius r_e has been found by Kormendy[13]:

$$\mu_B(r_e) = \left[3.02 \log\left(\frac{r_e}{1\ \text{kpc}}\right) + 19.74 - 3.02 \log(h)\right] B \text{ mag arc s}^{-2} \tag{3}$$

Combining this with the theoretical relationship between the absolute magnitude of a de Vaucouleurs galaxy, r_e and $\mu(r_e)$:

$$M_B = -39.96 + (\mu_B(r_e)/1 \text{ mag arc s}^{-2}) - 5 \log\left(\frac{r_e}{1\ \text{kpc}}\right) \tag{4}$$

gives $L_B = L_{B,K} h^{x_1} (r_e/1\ \text{kpc})^{x_2}$, where the empirically determined constants are $x_1 = 1.208$, $x_2 = 0.792$ and $L_{B,K} = 9.5 \times 10^{35}$ W.

A second empirical relationship found for elliptical galaxies is the $L - \sigma^4$ correlation, connecting luminosity with observed central line-of-sight velocity dispersion. Using the particular form found by Terlevich *et al.*[14]:

$$\log\left(\frac{\sigma}{1 \text{ km s}^{-1}}\right) = -0.0959 M_B + 0.328 + 0.4795 \log(h) \tag{5}$$

we find $L_B = L_{B,T} h^{-2} (\sigma/100 \text{ km s}^{-1})^{x_3}$, where $x_3 = 4.171$ and $L_{B,T} = 7.3 \times 10^{34}$ W. Finally we use the theoretical relation between σ and M_T/r_e, given[15] for small finite apertures by

$$\sigma^2 \simeq 0.2\, G M_T / r_e \tag{6}$$

Expressing B-luminosities in terms of the characteristic luminosity L_B^* which occurs in luminosity functions of the form proposed by Schechter[16], and defining $L_B^* = L_{B,n}^* h^{-2}$, we thus obtain

$$M_T(l) = 1.2 \times 10^{10} \left(\frac{L_{B,n}^{*x_5}}{L_{B,T}^{2/x_3} L_{B,k}^{1/x_2}}\right) h^{-(x_4 + 2x_5)} l^{x_5} M_\odot \tag{7}$$

where $x_4 = x_1/x_2 - 4/x_3 = 0.566$, $x_5 = 1/x_2 + 2/x_3 = 1.742$ and $l = L_B/L_B^*$. This implies (using our adopted parameter values) $M_T/L_B \propto M_T^{0.43}$ in good agreement with data presented[17] by de Vaucouleurs.

Combining these relations, the wind criterion equation (2) can be expressed directly in terms of the present value of the dimensionless luminosity l, cosmological epoch t, and other observable quantities. We find, for outflow beyond radius x

$$C \equiv C_0 h^{(2+x_1)/x_2} l^{-(1/x_2 + 4/x_3 - 1)} \frac{t}{t_0} \geqslant F_{\text{crit}}(x) \tag{8}$$

where t_0 is the present epoch, and the constant C_0 depends entirely on presently observed quantities such as t_0, $\alpha(t_0)$, $\alpha_{sn} E_{sn}(t_0)$, $L_{B,K}$, $L_{B,T}$, $L_{B,n}^*$ and the ratio L_T/L_B of total-to-blue luminosity. (This arises from the definition[11] of ν_{sn}.) Adopting for this last factor the reasonable value of 10, and taking[18] $L_{B,n}^*$ to correspond to an absolute blue magnitude $M_{B,n}^* = -21.0$, we obtain $C_0 \simeq 1.5$, showing that this coefficient is of the order of unity. The value of C_0, although uncertain, can in principle be determined entirely by independent observations; the time-dependence shown in equation (8) comes from the assumed time-dependence of $\alpha_{sn} E_{sn}/\alpha$ as discussed above.

Inverting equation (8) we find that at epoch $\tau = t/t_0$ all galaxies with luminosity exceeding $l_c(\tau)$ will have non-zero stagnation radii, where

$$l_c(\tau) = (C_0/F_{\text{crit}}(0))^{x_6} h^{x_7} \tau^{x_6} \tag{9}$$

and $x_6 = 1/(1/x_2 + 4/x_3 - 1) = 0.819$ and $x_7 = (2 + x_1) x_6 / x_2 = 3.316$.

The luminosity function for elliptical galaxies is similar to that of the general field[18]. We therefore assume a Schechter luminosity function defined so that the number of galaxies per unit volume with luminosities in the range $(L, L + \mathrm{d}L)$ is

$$\phi(L)\mathrm{d}L = (\phi^*/L^*)(L/L^*)^a \exp(-L/L^*)\mathrm{d}L$$

The comoving density of galaxies with non-zero stagnation radii normalized to the present epoch $\tau = 1$ is thus given by

$$D(\tau) = n(\tau)/n(1) \tag{10}$$

where

$$n(\tau) = \phi^* \int_{l_c(\tau)}^{\infty} l^a \exp(-l)\, \mathrm{d}l \tag{11}$$

On our model, in which inflowing stellar mass loss fuels active galactic nuclei, we expect the number density of radio sources at any epoch to be proportional to $n(\tau)$. Thus $D(\tau)$ also describes the comoving density of active galaxies, D_{act}.

Observationally it is found[19] that the comoving density of quasars is well fitted by an exponential in look-back time $T = 1 - \tau$, given by $D_{\text{obs}} \simeq \exp(13T)$. Figure 2 shows these observations compared with the prediction of our model using parameters $C_0 = 1.5$ (see above), $H_0 = 85$ km s^{-1} Mpc^{-1} (refs 20, 21) and luminosity function exponent $a = -1.10$ (ref. 18). The predicted curve is insensitive to choice of a, but is more critically determined by C_0, H_0 and $F_{\text{crit}}(0)$ through the relation (equation (9)) $l_c(1) \simeq 14$. An equally good fit to the data is obtained for other values of C_0 and H_0 provided $l_c(1)$ is approximately unchanged; for example good fits are also obtained with $(C_0, h) = (0.7, 2), (11.0, 1)$.

Assuming that inflowing stellar mass loss produces nuclear activity with luminosity given by $P = \eta c^2 \alpha(t) M_*(r_s(t))$ (for example, by accretion from a disk onto a massive black hole), we can use the present theory to predict the comoving luminosity function of active nuclei at various epochs. Figure 3 shows the results of such a calculation, made using $\phi^* = 2.35 \times 10^{-3}$ h^3 galaxies (Mpc)$^{-3}$ and energy generation efficiency factor $\eta = 0.3$. (The shapes of the comoving function do not depend on either ϕ^* or η.) The predicted luminosity function is in encouraging agreement with that indicated by observations[19], although a detailed comparison is not possible due to the lack of spectral information in such a simple model.

Finally, we have also calculated the numbers of galaxies with inflow as a function of redshift z for an open $q_0 = 0.05$ model

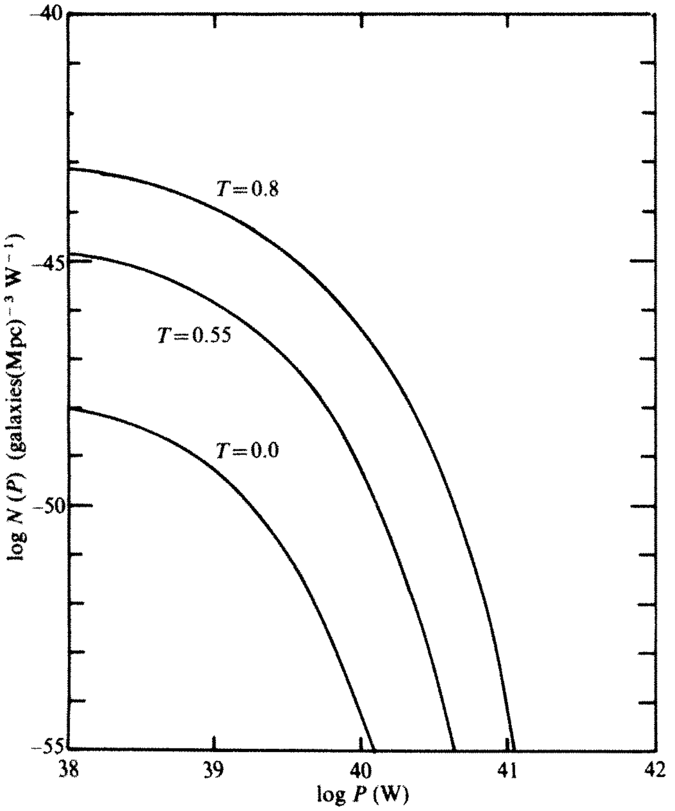


Fig. 3 The predicted co-moving bolometric luminosity function for an accreting black hole model, for three values of the look-back time T.

universe. The predicted counts can be fitted approximately by a powerlaw $N(<z) \simeq 1.5 \times 10^4 z^{5.85}$. The total number of sources out to redshift $z = 4$ is $\sim 5 \times 10^7$, consistent with counts of faint objects[22].

We conclude that the present theory apparently explains the large increase in comoving density of radio sources towards early epochs. At present only massive elliptical galaxies ($L \geq 10L^*$) should be radio sources. At earlier times, because of the higher stellar mass loss rate, less massive ellipticals cannot sustain total outflow. The absence of gas in moderately massive ellipticals may thus be a relatively recent phenomenon. On this model the reason for a rapid increase in comoving density of radio sources (by factors $\geq 10^3$ over a Hubble time) is principally the steepness of the luminosity function at $L > L^*$. This means that small evolutionary changes in l_c greatly affect the source counts. Whatever the nature of the central source of activity, the present work supports the hypothesis that nuclear activity in elliptical galaxies is caused by inflowing stellar mass loss.

We thank Professor R. J. Tayler and Dr R. C. Smith for helpful comments. This research was supported by the SRC.

Received 17 July; accepted 12 December 1980.

1. Johnson, H. E. & Axford, W. I. *Astrophys. J.* **165,** 381–390 (1971).
2. Mathews, W. G. & Baker, J. C. *Astrophys. J.* **170,** 241–259 (1971).
3. Bregman, J. N. *Astrophys. J.* **224,** 768–781 (1978).
4. Bailey, M. E. *Mon. Not. R. astr. Soc.* **191,** 195–206 (1980).
5. MacDonald, J. & Bailey, M. E. (in preparation).
6. de Vaucouleurs, G. in *Handbuch der Physik* Vol. 53 (ed. Flugge, S.) 311 (1959).
7. Gisler, G. R. *Astr. Astrophys.* **51,** 137–150 (1976).
8. Oemler, A. & Tinsley, B. M. *Astr. J.* **84,** 985–992 (1979).
9. Shklovsky, I. S. *Soviet Astr.* **22,** 413–419 (1978).
10. Whelan, J. & Iben, I. *Astrophys. J.* **186,** 1007–1014 (1973).
11. Tammann, G. A. *Astrophys. Space Sci. Lib.* **45,** 125–130 (1974).
12. Gorenstein, P. & Tucker, W. H. *A. Rev. Astr. Astrophys.* **14,** 373–416 (1976).
13. Kormendy, J. *Astrophys. J.* **218,** 333–346 (1977).
14. Terlevich, R., Davies, R. L., Faber, S. M. & Burstein, D. Preprint (1980).
15. Bailey, M. E. & MacDonald, J. *Mon. Not. R. astr. Soc.* **194,** 195–200 (1981).
16. Schechter, P. L. *Astrophys. J.* **203,** 297–306 (1976).
17. de Vaucouleurs, G. *IAU Symp.* **58,** 1–53 (1974).
18. Kirshner, R. P., Oemler, A. & Schechter, P. L. *Astr. J.* **84,** 951–959 (1979).
19. Turner, E. L. *Astrophys. J.* **231,** 645–652 (1979).
20. de Vaucouleurs, G. *Nature* **266,** 126–129 (1977).
21. Hanes, D. A. *Mon. Not. R. astr. Soc.* **188,** 901–909 (1979).
22. Bohuski, T. J. & Weedman, D. W. *Astrophys. J.* **231,** 653–658 (1979).

Thermal-mechanical properties of a monomolecular film detected by a vibrating reed

R. H. Lacombe & B. M. J. Kellner

International Business Machines Corporation, East Fishkill Facility, Hopewell Junction, New York 12533

Thin films of organic materials are finding many applications in modern technology. The microelectronics industry, in particular, uses very thin organic films both as photoresists and as insulating layers. The ultimate in organic thin films is represented by molecular monolayers of the order of 2.0 nm thick as deposited by the Langmuir–Blodgett technique[1]. By adapting a vibrating reed apparatus[2] we have investigated the thermal–mechanical properties of a supported 6-nm thick monolayer film. Although a monolayer melting transition could not be found in the resultant spectra, at least two other, low-temperature peaks (at −150 and −20 °C) were detected which seem to be due to relaxation processes both within the film as well as at interfaces of the monolayer structure. Our main aim was to determine whether the mechanical properties of ultrathin films could be measured by the vibrating reed technique, and to detect any thermal–mechanical relaxations which may occur in such films, in particular whether or not a melting transition could be detected at elevated temperatures. Such thermal–mechanical measurements have long been used in the plastics industry to characterize the mechanical properties of resin materials in bulk. The vibrating reed technique may provide an analogous service for very thin organic coatings.

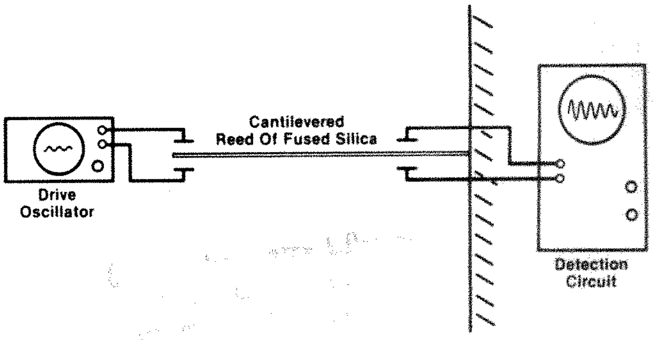


Fig. 1 The vibrating reed apparatus. The silica reed is used as a substrate to support thin films whose mechanical properties are to be investigated.

Other recent work in thin film mechanical spectroscopy has investigated freely suspended liquid crystal films using a disk and support ring apparatus[3–5]. Also the superfluid transition of ^{4}He has been studied using mylar as a substrate[6]. The vibrating reed technique supplements these methods by covering a wide range of temperature and frequency and by not being restricted to monolayer forming films.

Figure 1 gives a schematic representation of the vibrating reed apparatus which is described in detail in ref. 2. The basic concept is: a cantilevered reed is excited into vibration by a set of drive electrodes connected to a suitable oscillator; a corresponding set of detection electrodes, positioned approximately one-third of the way up from the base (a position which avoids setting the electrodes at a vibration node) is used to monitor the amplitude of vibration.

First the reed is excited to a given amplitude and then the drive signal is switched off and the free decay of the reed vibration is monitored. Theoretically, there is an infinite number of modes of vibration into which the reed can be excited by appropriately adjusting the drive frequency. In practice, one works with the first dozen or so lowest tones. The dynamic modulus of the reed material is directly related to the frequency of vibration at any given tone, and the internal friction of the material is inversely proportional to the number of cycles taken for the amplitude to decay to one-half of its initial value. The precise details on measuring these quantities are given in ref. 2.

In the present experiment, a composite sample is used. It consists of a reed-like substrate of fused silica, ~50 μm thick. When operating under a vacuum $\sim 10^{-6}$ torr, the internal friction of the fused silica substrate is exceedingly low over a wide temperature range. This characteristic allows the internal friction of coatings to be detected which may be even 1,000 times thinner than the substrate itself.

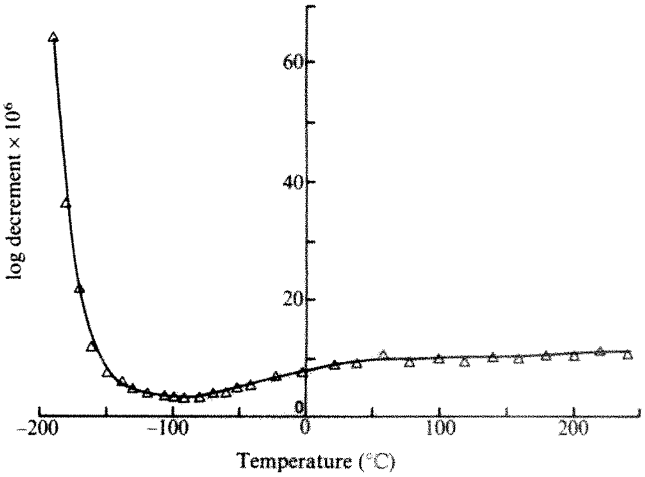


Fig. 2 Plot of internal friction of fused silica from −190 to 240 °C. Note the low and slowly varying behaviour of the internal friction above −150 °C. The solid line was obtained through a piecewise least-square polynomial fit to the data.

0028-0836/81/080661—02$01.00

Figure 2 shows the internal friction of fused silica from liquid nitrogen temperature to 240 °C. Note the relatively flat and low internal friction behaviour above −150 °C. Below −150 °C, fused silica shows a large internal friction peak[7] which would make it an unsuitable substrate material in this temperature range. Above −150 °C, however, the internal friction is around 10^{-5}, which implies that typically 10^5 cycles are required for the blank reed to decay to half of its initial amplitude. The solid line in Fig. 2 was used to interpolate the experimental data so that the contribution of the bare substrate to the internal friction of the composite sample could be subtracted accurately and consistently.

An automated Langmuir trough was modified for Langmuir–Blodgett deposition[8]. Constant pressure conditions during dipping were maintained by an electric feedback circuit between the surface pressure monitor and the drive for the movable barrier of the trough. Arachidic acid was spread on an aqueous substrate of 3×10^{-4} M cadmium chloride buffered with 5×10^{-4} M sodium acetate and adjusted to *p*H 6 by the addition of small amounts of acetic acid and/or sodium hydroxide. The film was then compressed to 30 mN m^{-2} and transferred by the Langmuir–Blodgett technique onto a fused silica reed which had been scrupulously cleaned. A film composed of three monolayers of cadmium arachidate[9] on each side of the reed was built up. A fatty acid salt deposited in this way has been shown[10] to be composed of a molecular array in which the alkane chains are close-packed and the metal ions are sandwiched between the carboxyl groups of adjacent monolayers.

Figure 3 summarizes the primary experimental results of this work showing the internal friction of the built-up monolayer film as a function of temperature. The data represent three separate runs where the sample was heated to progressively higher temperatures to see if a melting transition could be detected. Such a transition in cadmium arachidate would be expected to occur at ~140 °C. This expectation is based on an extrapolation from the electron diffraction and frictional data reported by Bowden and Tabor[11]. The high-temperature data, however, showed only a flat baseline behaviour similar to that for the bare reed (see Fig. 2).

We conclude that material from the upper layers of the three-layer film simply sublimed away at the higher temperatures. Therefore no phase transition to a liquid form was observed. However, Fig. 3 clearly shows that there are at least two distinct dispersion peaks at the lower temperatures −150 and −20 °C, respectively. The dramatic decrease in the amplitude of the peak near −150 °C after successive heatings to higher and higher temperatures would seem to indicate that this peak is due to relaxation processes occurring in the bulk of the composite film. The higher temperature peak near −20 °C, however, shows a significantly less decrease in amplitude due to heat treatment. This peak may in fact be due to relaxations occurring at the interfaces of the monolayer structure (either monolayer–monolayer or monolayer–silica). This would also be consistent with the much lower intensity of this peak compared with the one near −150 °C. Figure 4 gives an expanded representation of the mechanical spectrum near −20 °C.

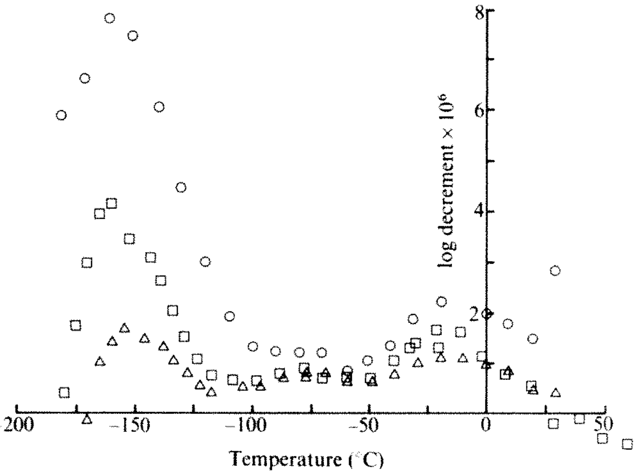


Fig. 3 Internal friction of three molecular layers of cadmium arachidate: ○, heated to 130 °C; □, heated to 160 °C; △, heated to 250 °C. Two dispersion peaks are clearly visible near −150 °C and −20 °C. This plot was obtained by subtracting the internal friction of the bare reed substrate from that of the coated reed. However, a correction factor should be included which depends on the modulus of the substrate and the film and also accounts for the difference in thickness of the two. As there was no way to obtain a reasonable estimate of the modulus of the monolayer film, the data are only proportional to the true internal friction of the film. (See ref. 2 for further details.)

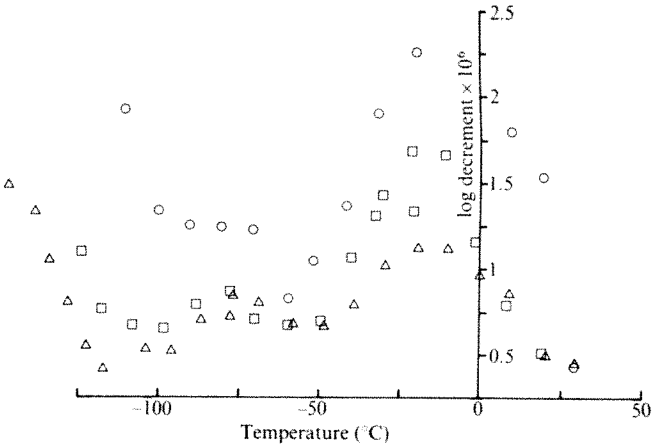


Fig. 4 Expanded view of Fig. 3 near −20 °C. A peak seems to be emerging near −75 °C for the samples which were heated to 160 and 250 °C.

The data in Figs 3 and 4 must be treated with caution as an exceedingly small amount of material is being dealt with, and there are no corroborative data on similar types of samples. However, these data may be clarified by comparison with another physical system which involves long hydrocarbon chains, for example, bulk polyethylene. Low density polyethylene also exhibits two low-temperature transitions near −100 °C and −10 °C respectively. The peak near −100 °C (referred to as the γ peak) seems to occur in polymers with five or more CH_2 units in series and is usually associated with amorphous or interfacial regions in semicrystalline systems. The peak near −10 °C (known as the β peak) is also associated with crystalline imperfections or possibly grain boundaries[12]. These observations support our conclusions if one remembers that our sample film thickness of 6 nm is about the size of an 'interfacial' region in a bulk polymer. Although these conclusions are preliminary, the fact that clear, reproducible spectra can be obtained indicates that the vibrating reed technique has a strong potential as an analytical tool for the study of the mechanical properties of ultrathin films.

We thank B. Berry and W. Pritchet for technical assistance and advice.

Received 10 September; accepted 9 December 1980.

1. Gaines, G. L. Jr in *Insoluble Monolayers at Liquid–Gas Interfaces* Ch. 8 (Wiley-Interscience, New York, 1966).
2. Berry, B. S. & Pritchet, W. C. *I.B.M. J. Res. Devl.* **19,** 334 (1975).
3. Pindak, R., Bishop, D. J., Osheroff, D. D. & Sprenger, W. O. *Paper int. Conf. on Ordering in Two Dimensions*, Lake Geneva, Wisconsin (1980).
4. Pindak, R., Bishop, D. J. & Sprenger, W. O. *Phys. Rev. Lett.* **44,** 1461 (1980).
5. Tarczon, J. C. & Miyano, K. *Paper int. Conf. on Ordering in Two Dimensions*, Lake Geneva, Wisconsin (1980).
6. Bishop, D. J. & Reppy, J. D. *Phys. Rev. Lett.* **40,** 1727 (1978).
7. Anderson, O. L. & Bömmel, H. E. *J. Am. ceram. Soc.* **38,** 125 (1955).
8. Kellner, B. M. J. *Tech. Rep. TR22.* 2326 (International Business Machines Corporation, Hopewell Junction, New York, 1980).
9. Anderson, H. R. Jr. & Swalen, J. D. *J. Adhes.* **9,** 197 (1979).
10. Stephens, J. F. *J. Colloid Interf. Sci.* **38,** 577 (1972).
11. Bowden, F. P. & Tabor, D. *The Friction and Lubrication of Solids*, Part I, 209 (Oxford University Press, 1964).
12. Arridge, R. G. C. *Mechanics of Polymers*, 95–96 (Clarendon, Oxford, 1975).

Grain boundary microchemistry and stress-corrosion failure of mild steel

E. D. Hondros & C. Lea

National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

A common cause of failure of large metallic components is stress corrosion, often in subtle combination with fatigue. In the case of carbon steel as used in pipelines, the problem of understanding the cause of variations in susceptibility to intergranular stress corrosion between steel batches of apparently similar origin, history and operating conditions has been researched for many years. We report here the results of investigations including Auger electron spectroscopy which show that the source of variable behaviour is probably the differences in the microchemical content of grain boundary paths, differences which may easily exist between one batch or another of apparently similar steel.

Pure iron is resistant to stress corrosion cracking in nitrate environments, and although attempts have been made to trace the effects of a large number of alloying element additions, such as carbon, aluminium, chromium and nickel, there is apparently little systematic behaviour reported[1,2]. Most previous studies have been carried out on steels containing the usual amounts of the normal individual impurities common to commercial steel stock. Our present investigations suggest that the lack of reproducibility in the reported effects could be due to differences in heat treatment and impurity contents.

Previous Auger electron spectroscopy measurements[3] established that different impurity elements may enrich grain boundaries and other interfaces by factors between 1 and 10^5. Thus, on the assumption (later confirmed experimentally) that impurity elements at low bulk levels do not affect mechanical properties such as yield stress, we undertook a direct study of the microchemistry of the grain boundaries to investigate whether a correlation between susceptibility and bulk impurity content exists.

The test specimens were prepared from 11 ingots of high purity 0.1% C mild steel (Table 1). In 10 of these alloys, the free residual sulphur was tied up chemically and rendered harmless by the addition of 0.05% Mn. To these ingots, a known, small quantity of one commonly occurring impurity element was added separately at that particular bulk level calculated to give rise to a grain boundary segregation content of about 20% of a monatomic layer following an isothermal heat treatment of 100 h at 600 °C. This treatment does not affect the microstructure nor the bulk properties but allows the 20% monolayer segregation to be attained from bulk impurity levels comparable with those encountered in commercial stock. The additions were: 0.3% Al, 0.15% As, 0.01% Ca, 0.1% Cu, 0.3% Ni, 0.03% P, 0.08% Sb, 0.1% Sn and 0.2% Zn. The amount of bulk additive sufficient in each case to produce this 20% monolayer segregation was calculated on the basis of a correlation against maximum solid solubility established earlier[1]. From this, the grain boundary enrichment ratio β_N in a dilute binary alloy is given, within the limits,

$$\beta_N = \frac{1 \text{ to } 10}{X_N} \qquad (1)$$

where X_N is the solid solubility limit of element N and β_N is the concentration of the solute in the boundary at equilibrium compared with that in the bulk.

While such quantities were required in advance as part of the experimental design, the predicted segregation levels were later confirmed by exposing the grain boundaries in the spectrometer to direct analysis. In most cases this was difficult, because the levels of impurities were so low that they did not induce sufficient low temperature brittleness to allow exposure of the boundaries by *in-situ* fracture. The successful procedure was to

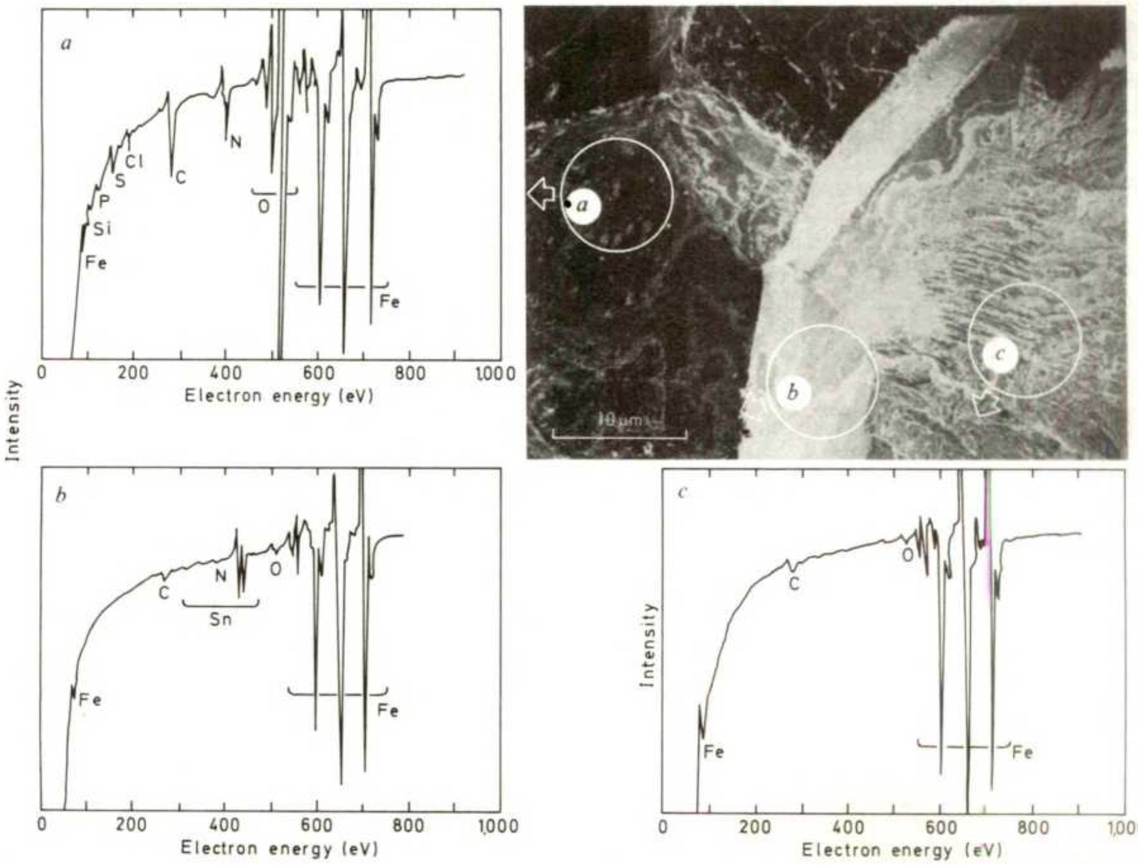


Fig. 1 Auger electron spectra taken from the three regions at a stress corrosion crack front: *a*, the intergranular corroded crack; *b*, the intergranular impact crack; and *c*, the transgranular impact crack.

0028-0836/81/080663—03$01.00

Table 1 Analytical results

Cast	Code	Analysed bulk impurity level wt%	Predicted grain boundary content after 100 h at 600 °C (X_B) (monolayers)	Measured grain boundary content (X_B) (monolayers)	Measured corrosive–inert ratio (R)	Relative increase for constant bulk level
1	(S)	0.0026	0.68*	0.80±0.25	0.40±0.05	700±140
2	(pure)	—	—.	—	0.86±0.06	—
3	(Al)	0.295	0.18	0.17±0.14	0.57±0.14	1.0±0.3
4	(As)	0.14	0.18	0.19±0.13	0.81±0.17	0.4±0.1
5	(Ca)	0.009	0.15*	0.09±0.07	0.66±0.24	27±11
6	(Cu)	0.10	0.18	0.14±0.05	0.67±0.05	1.9±0.3
7	(Ni)	0.305	0.20	0.10±0.3	0.79±0.12	0.2±0.05
8	(P)	0.028	0.18	0.27±0.13	0.38±0.07	20±5
9	(Sb)	0.076	0.17	0.20±0.07	0.80±0.11	0.9±0.2
10	(Sn)	0.096	0.18	0.26±0.04	0.78±0.09	1.0±0.2
11	(Zn)	0.169	0.15	0.16±0.06	0.81±0.17	0.3±0.1

* Not at equilibrium.

arrest the stress corrosion tensile test before rupture and to transfer the sample from the corrosive solution into the spectrometer, where the fracture was completed by low temperature impact in the ultra-high vacuum. By this process, the pre-existing intergranular stress corrosion crack sometimes extended up to one grain into the virgin steel before the fracture continued in a cleavage mode. These freshly exposed grain boundaries could be isolated and analysed, yielding the actual quantity of segregant on the grain boundaries before stress corrosion failure. Figure 1*a* shows that the stress corrosion crack was heavily oxidized and contaminated by exposure to air and the vacuum environment of CO. Figure 1*b* shows the intergranular fracture surface exposed within the spectrometer and immediately analysed, revealing here a tin segregation and Fig. 1*c*, the remainder of the fracture, which is transgranular and hence shows essentially only iron. The carbon Auger electron signal is associated with monolayer adsorption of CO from the vacuum environment and not with bulk carbon. This is demonstrated by the removal characteristic during ion sputtering of the fracture surface. Using quantification procedures for submonatomic layers, the segregation levels of the added impurities are expressed in terms of fractional adsorbed monolayers, X_B, on the grain boundaries. Table 1 confirms the ~20% level predicted by application of equation (1), within the limits of the equation.

The samples were hot rolled into standard specimen blanks and heat treated in air at 800 °C for 2 h, followed by the isothermal segregation treatment. Flat, unnotched tensile specimens were prepared and tested to failure in an environmental cell of flowing NH_4NO_3 (5M) solution with parallel control tests in a stable environment of liquid paraffin at 80 °C, all at a constant strain rate of $2\times10^{-6}\ s^{-1}$. A similar set of experiments were carried out in constant load testing conditions on a four-point bend rig, submerged in a static bath of the same solution.

The results of the constant strain rate tests are shown in Fig. 2. The error bars in the plotting values indicate the spread in data over six to eight measurements. Every specimen type exhibited stress corrosion in the nitrate environment, failing in a time shorter than that in the control environment. The additions, Zn, As, Sb, Ni and Sn, have a minor effect in reducing rupture lives, Cu, Ca and Al have a modest effect, while S and P clearly have the greatest influence. In the nitrate solution, fracture was predominantly and often entirely, intergranular, while in the control solution, ductile fracture occurred. Table 1 shows the average times to failure in the nitrate solution compared with those in the control solution (corrosive–inert ratio, R). Not only are the yield stresses of all the samples within the range $220\ MN\ m^{-2}\pm5\%$, but Fig. 2 shows that the minor impurity additions (as reflected in the times to failure in the inert environment), have not introduced any significant secondary effects on mechanical properties.

We present the results in terms of the change in R, in the analytical form:

$$-\Delta R = 20\mathrm{P} + 1.9\mathrm{Cu} + 1.0\mathrm{Sn} + 0.9\mathrm{Sb} + 0.4\mathrm{As} + 0.3\mathrm{Zn} + 0.2\mathrm{Ni} + (700\mathrm{S} + 27\mathrm{Ca} + 1\mathrm{Al}) \quad (2)$$

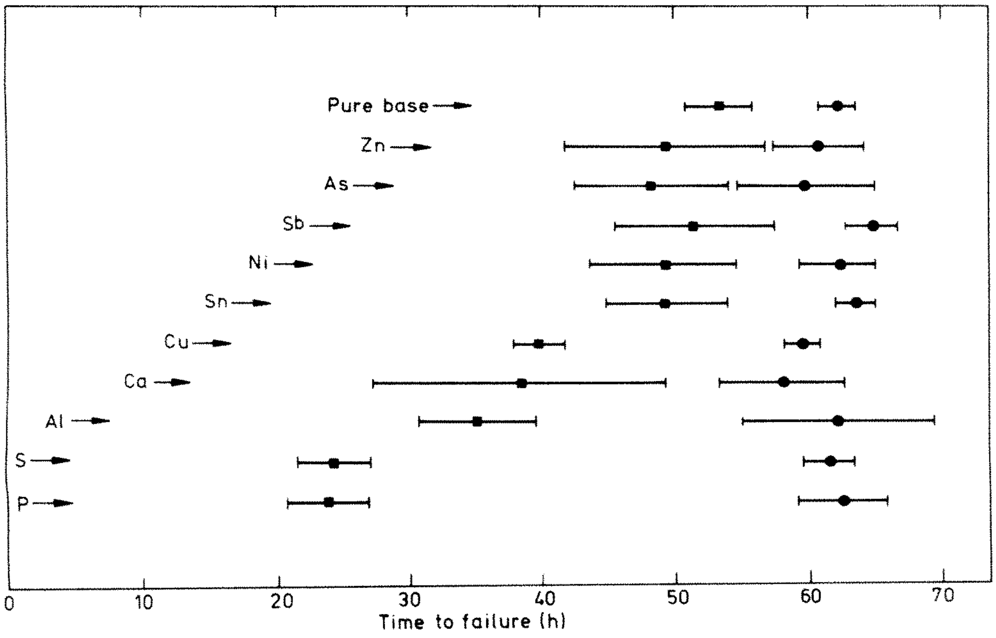


Fig. 2 The fracture times of the 11 alloys stress corroded at a constant strain rate. ●, Tests in paraffin; ■, tests in NH_4NO_3 solution.

Here, the element symbols represent the bulk content in weight per cent. By introducing the measured enrichment ratio β_N, the changes in R are presented in terms of the bulk concentration of the elements, thus providing a ranking order of potent segregating impurities. In commercial steels, S, Ca and Al are mopped up as precipitates and do not have a segregation role. Thus, to optimize the performance of carbon steels in nitrate environments, ΔR must be minimized. This can be achieved by reducing the bulk impurity levels of the first few elements, but in particular, that of residual phosphorus.

The detrimental influence of phosphorus on the stress corrosion properties of mild steel has been shown by Rädeker and Mishra[4] who added phosphorus in the range 0.012–0.12% to a pure base stock. By using the free energy of segregation obtained from our data, these bulk P levels can be readily transposed to grain boundary segregation levels and the correlation of these X_B values with stress corrosion properties is very similar to that reported here.

We have tried to rationalize the present results in terms of a mechanism whereby a galvanic potential is sustained and the grain boundary is exposed by preferential dissolution of atoms from the boundary. In an oxidizing environment such as aqueous ammonium nitrate, the driving force for atomic dissolution is related to the potential difference between the matrix and the segregants forming a galvanic cell. The hierarchy of the measured differences in the corrosive–inert ratio (ΔR) between each alloy and that of the pure base, normalized to a grain boundary concentration of 1% monolayer, bears a close resemblance to the order of the differences ΔE_0 between the equilibrium potentials of each segregant and that of iron. The potency of each added element for promoting stress corrosion cracking thus increases as $|\Delta E_0|$ increases. The coefficients of other elements, not shown in Fig. 2 or equation (2) can, therefore be predicted approximately as the product of the difference in equilibrium potential from that of iron and the grain boundary enrichment of the element[5]. Thus, although Al, Ti and Ca, for example, all have large values of ΔE_0 they do not segregate, their β values are low, and are therefore not harmful. (Indeed their addition may be beneficial by scavenging those elements which do readily segregate.) Thus, the application of the difference in equilibrium potential plus the grain boundary enrichment factor, seems an adequate first test for sensitivity of a particular element in intergranular stress corrosion.

Received 19 September; accepted 19 December 1980.

1. Parkins, R. N. *J. Iron Steel Inst.* **172,** 149 (1952).
2. Uhlig, H., Perumal, K. E. & Talerman, M. *Corrosion* **30,** 229 (1973).
3. Seah, M. P. & Hondros, E. D. *Proc. R. Soc.* A**335,** 191 (1973).
4. Rädeker, W. & Mishra, B. N. *Werkstoffe Korros.* **21,** 691 (1970).
5. Lea, C., Seah, M. P. & Hondros, E. D. *Mater. Sci. Engng* **42,** 233 (1980).

Controls on low-stress hydro-fracture dilatancy in thrust, wrench and normal fault terrains

Richard H. Sibson

Department of Geology, Imperial College, London SW7 2BP, UK

Arrays of parallel extension fractures and veins are sometimes associated with exhumed faults and seem to be the product of repeated hydro-fracturing under a shared stress regime. Here theoretical considerations are used to suggest that this type of hydro-fracture dilatancy is essentially a low differential stress phenomenon. It may develop under hydrostatic fluid pressures around normal faults at shallow depths, but can only occur in association with thrusts when fluid pressures exceed the lithostatic load. This contrasts with most existing models for dilatancy, where the effects only become pronounced at high stress levels and are best developed around thrust faults.

Pre-failure dilatancy, leading to the redistribution of fluids, is a possible cause of observed precursors of shallow earthquakes[1,2], although the extent to which fluid migration is necessarily involved has been questioned[3]. Both the size of the dilatant zone and its stress-dependence are affected by the various types of dilatancy which can occur[4]. However, dilatant effects are commonly attributed to the opening of microcracks under high differential stresses. This accords with the more widespread recognition of earthquake precursors in association with thrust faults, where high stress levels may be characteristically expected[5] and subhorizontal extension cracks may open vertically without restraint[6].

The origin of an observable form of macroscopic fracture dilatancy which is sometimes found around ancient, exhumed faults (Figs 1 and 2*a*) is considered here. The arrays of parallel extension fractures or veins bear a systematic relationship to the faults in accordance with their slip-sense, which suggests that they formed in the same stress field. For example, in Fig. 1, vertical extension veins infilled with fibrous gypsum are associated with parallel-striking normal faults which disrupt a flat-lying sequence of Neogene shales and sandstones at Konarak in south-east Iran[7]. Both sets of structures are compatible with a regime where the maximum principal compressive stress, σ_1, is vertical and the least compressive stress, σ_3, is horizontal and perpendicular to strike. Where such extensional veins abut onto faults they are known as 'feather veins'[8,9]. The vein-fills usually consist of hydrothermal minerals which have grown inwards as fibres from the walls towards a median parting, the texture characteristic of a progressively opening fissure[10]. In many cases, a history of incremental opening is apparent[11], the implication being that fluid flow occurred intermittently with each episode terminating as deposition of hydrothermal minerals destroyed fracture permeability. Self-sealing of fractures in this manner is an important process controlling permeability in active hydrothermal systems[12].

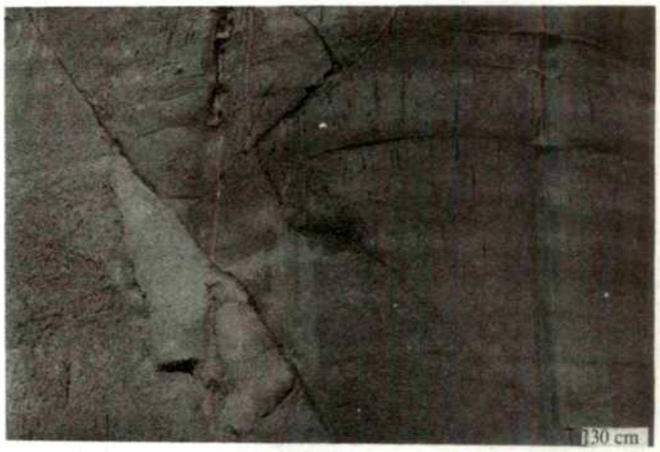


Fig. 1 Conjugate normal faults associated with vertical, gypsum-filled extension veins, Konarak, Makran Coast of south-east Iran. Clastic sandstone dykes (one occurs by the hammer) cut across the gypsum veins before feeding onto the faults.

These textural characteristics, coupled with the typically high length to thickness ratios, suggest that such veins are the product of repeated hydraulic extension fracturing[13]. Similarly oriented barren extension joints may also have formed by hydro-fracturing, but do not retain evidence for more than one episode of opening. Deposition of hydrothermal minerals ensures that some permanent dilatant strain results; at Konarak this locally exceeds 10^{-2} close in to the faults but the veining dies out 100 m or so to each side. From the presence of gypsum, implying temperatures <50 °C (ref. 14), and the observation that at least one of the normal faults displaces overlying Quaternary raised

0028–0836/81/080665—03$01.00

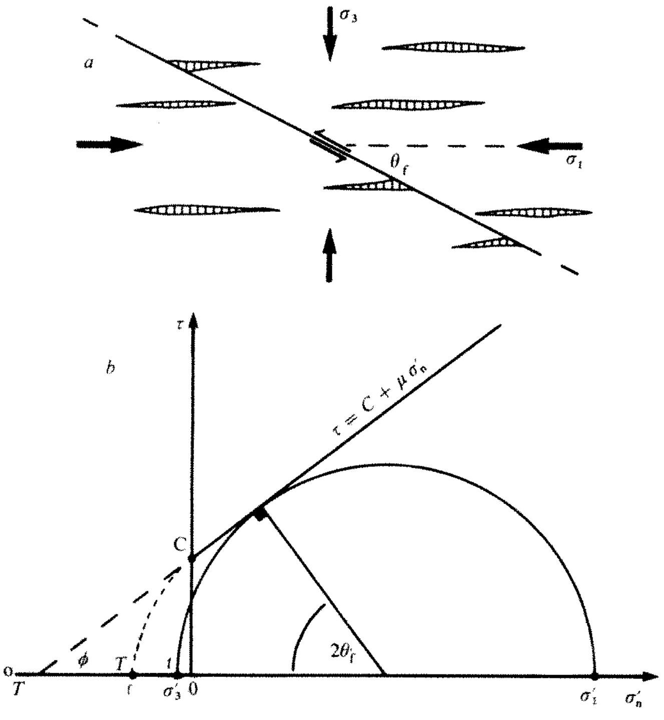


Fig. 2 Hydraulic extension fracturing in association with faults: *a*, field relationships; *b*, critical stress condition for simultaneous shear failure and reopening of extension fractures.

beaches[7], the vein arrays at Konarak have clearly developed in near-surface conditions. However, similar arrays with quartz or calcite as the usual vein filling may occur throughout the frictional regime occupying perhaps the top 10–15 km of quartzo-feldspathic crust[15]. Indeed, Ramsay[11] has attributed extension veins showing multiple increments of opening and sealing (>500 in one case) to repeated hydro-fracturing in upper greenschist facies conditions.

To examine the conditions in which cyclic hydro-fracture dilatancy can occur adjacent to an intermittently slipping fault, the case of an existing fault and extension fracture array in otherwise homogeneous rock is considered (Fig. 2*a*). The simple law of effective stress[16] is assumed to hold so that effective normal stresses are given by

$$\sigma_n' = \sigma_n - P \qquad (1)$$

where σ_n is the applied normal stress and P is the pore fluid pressure. Effective principal compressive stresses are thus $\sigma_1' > \sigma_2' > \sigma_3'$, and failure criteria can be represented on a Mohr plot of shear stress, τ, versus σ_n' (Fig. 2*b*). In the compressional field, shear failure along the fault is taken to be governed by a frictional criterion of Coulomb form,

$$\tau = C + \mu\sigma_n' \qquad (2)$$

where C is the cohesive strength of the fault and the coefficient of friction $\mu = \tan\phi$, ϕ being the angle of friction. The fault is also assumed to be oriented at the optimum angle to σ_1 for frictional sliding, so that $\theta_f = 45° - \phi/2$. Equation (2) may then be rewritten in terms of the effective principal stresses at failure[15] giving

$$\sigma_1' = 2K^{\frac{1}{2}}C + K\sigma_3' \qquad (3)$$

where $K = ((1+\mu^2)^{1/2} + \mu)^2$. A reasonable value[5] of $\mu = 0.75$ gives $\phi = 37°$, $\theta_f = 27°$ and $K = 4$. The extension fractures or veins are assumed to have a 'healed' tensile strength, t, which is less than the tensile strength of the intact rock, T, and is also small in comparison with C (strictly, $t < \frac{1}{2}C$ for $\mu = 0.75$), so that shear failure is restricted to the compressional field. Thus the condition for reopening the extension fractures is

$$\sigma_3' = -t \quad \text{or} \quad P = \sigma_3 + t \qquad (4)$$

Hydraulic fracturing, requiring $\sigma_3' < 0$, can only occur when $C > 0$. Thus, hydro-fracturing should not develop adjacent to cohesionless faults. However, natural faults are likely to possess a small cohesive strength through a combination of surface roughness and hydrothermal cementation. Its value is unlikely to exceed 10 bars (ref. 17) unless cementation is marked, in which case it may approach the cohesive strength of intact rock, perhaps 100 bar. If the opening of hydraulic fractures accompanies faulting, an important constraint is placed on the prevailing level of differential stress. From Fig. 2*b* and equations (3) and (4), the critical differential stress for simultaneous shear failure and hydro-fracturing is

$$\begin{aligned}(\sigma_1 - \sigma_3) &= 2(C - \mu t)((1+\mu^2)^{1/2} + \mu) \qquad (5)\\ &= 4C - 3t \text{ for } \mu = 0.75\end{aligned}$$

Above this critical stress hydro-fracturing is impossible and only shear failure can occur. Accepting $C = 10$ bar, and taking the extreme case when $t \to 0$, the maximum differential stress allowing hydro-fracturing is just 40 bar. It will be correspondingly greater for intact rocks or a hydrothermally cemented fault, but in comparison with the differential stresses of several kilobars generally required to induce microcrack dilatancy[18], hydro-fracture dilatancy is clearly a low-stress phenomenon.

This critical stress condition may be considered in relation to Anderson's[19] three modes of faulting (based on whether the vertical stress $\sigma_v = \sigma_1$, σ_2, or σ_3), and the ratio of fluid to overburden pressure, $\lambda_v = P/\sigma_v = P/\rho g z$ (ρ being the average rock density, g the gravitational acceleration and z the depth). Previous studies[13,20] have shown that λ_v should be respectively greater and less than unity for hydro-fracturing in thrust and normal fault terrains, but have not taken full account of this limit to differential stress.

For thrust faults $\sigma_v = \rho g z = \sigma_3$, so from equation (4) the condition for hydro-fracturing is

$$\lambda_v = 1 + t/\rho g z \qquad (6)$$

requiring $\lambda_v \geqslant 1$ always. In the case of simultaneous hydro-fracturing and normal faulting, $\sigma_v' = \rho g z\ (1 - \lambda_v) = \sigma_1'$, so from equations (3) and (4),

$$\rho g z (1 - \lambda_v) = 2K^{\frac{1}{2}}C - Kt \qquad (7)$$

For simultaneous wrench faulting and hydro-fracturing, the particular situation where $\sigma_v' = \sigma_2' = \frac{1}{2}(\sigma_1' + \sigma_3')$ is considered. Then again from equations (3) and (4),

$$\rho g z (1 - \lambda_v) = K^{\frac{1}{2}}C - \tfrac{1}{2}t(K+1) \qquad (8)$$

Equations (6), (7) and (8) can be used to separate fields of potential shear failure and hydro-fracturing for the three faulting modes on plots of λ_v versus depth (Fig. 3). As an example, the delimiting curves have been drawn for the top 10 km of a crust with average density $\rho = 2.55$ g cm^{-3}, containing a fault–fracture system with failure parameters, $C = 50$ bar, $t = 10$ bar and $\mu = 0.75$. The barbed side of each curve represents the field of potential hydro-fracturing, except for the thrust regime where the curve marks conditions in which hydro-fracturing must occur. Note that the curve for wrench faulting is displaced towards those for normal and thrust faulting as the value of σ_2 moves respectively towards σ_1 and σ_3 from the median case considered.

Fluid pressures at least equivalent to the hydrostatic head ($\lambda_v = 0.39$ in this case) may generally be expected in established fault zones[15], so that in the uppermost regions of both normal and wrench faults, hydro-fracturing should inevitably accompany faulting. This may account for the extensive fissuring often observed in association with normal faults at the surface[21]. For hydro-fracturing at greater depths, fluid pressures must exceed

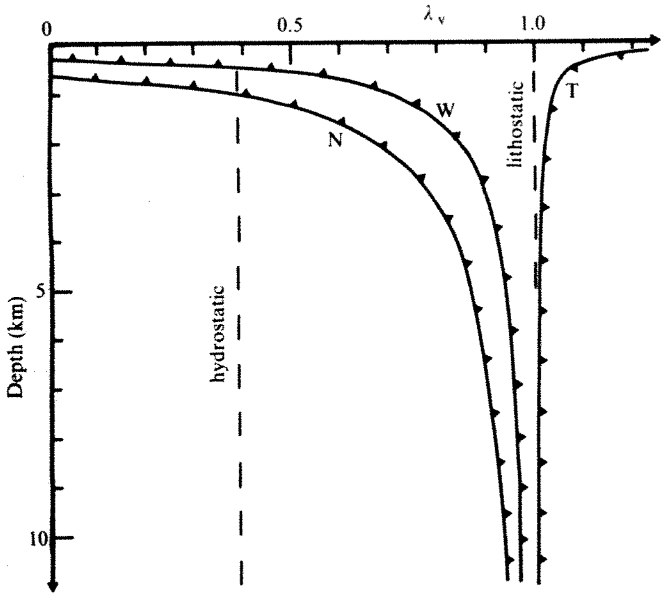


Fig. 3 Curves delimiting potential fields of shear failure and hydro-fracturing (barbed side) for the different faulting modes (T, thrust; W, wrench; N, normal). ($\rho = 2.55$ g cm^{-3}, $\mu = 0.75$, $C =$ 50 bar and $t = 10$ bar).

the hydrostatic head for all faulting modes. Such abnormal fluid pressures may arise through a variety of mechanisms but the presence of highly impermeable 'cap' layers is generally necessary[22]. However, throughout the frictional regime the required λ_v values remain significantly lower for normal faults than for thrusts, with those for wrench faults lying between. I have found that arrays of vertical extension veins around normal faults, as at Konarak, tend to be more common than subhorizontal veins associated with thrust faults. This contrasts directly with the common view[6] that dilatant effects should be more pronounced around thrust faults. Hydro-fracturing (particularly the opening of vertical extension fractures in normal and wrench fault terrains) may clearly also act as a regulatory mechanism, preventing fluid pressure from rising above the critical λ_v value for a given depth and faulting mode, once it exceeds the hydrostatic head.

The relationship between the onset of hydro-fracture dilatancy and the earthquake stress cycle is uncertain, and may vary with faulting mode. However, at Konarak there is evidence for the partial collapse of the dilatant vein array at failure in the form of clastic sandstone dykes which have been mobilized to cut across the gypsum extension veins before feeding onto the faults (Fig. 1). For pre-slip hydro-fracturing during stress accumulation at other than shallow depths around normal and wrench faults, fluid pressure must be sufficiently high for λ_v to exceed or attain the critical value for the particular depth, faulting mode and failure parameters (Fig. 3). Once hydro-fracturing has occurred, fluid pressures should drop towards the hydrostatic, resulting in an increase in effective normal stress and the strengthening of the fault. Shear failure then becomes possible once differential stress has risen sufficiently to overcome this dilatancy hardening effect.

I thank Dr B. K. Atkinson and Dr N. J. Price for helpful comments. Fieldwork at Konarak was accomplished during the 1976 Royal Geographical Society–Imperial College Expedition ot the Iranian Makran. I also thank the expedition members for assistance.

Received 22 September; accepted 8 December 1980.

1. Nur, A. *Bull. seism. Soc. Am.* **62,** 1217–1222 (1972).
2. Scholz, C. H., Sykes, L. R. & Aggarwal, Y. P. *Science* **181,** 803–810 (1973).
3. Mjachkin, V. I., Brace, W. F., Sobolev, G. A. & Dieterich, J. H. *Pure appl. Geophys.* **113,** 169–181 (1975).
4. Nur, A. *Pure appl. Geophys.* **133,** 197–206 (1975).
5. Sibson, R. H. *Nature* **249,** 542–544 (1974).
6. Nur, A., Bell, M. L. & Talwani, P. *Proc. Conf. on Tectonic Problems of the San Andreas Fault System* Vol. 13, 391–404 (Stanford University, 1973).
7. Vita-Finzi, C. *Geol. Min. Surv. Iran Rep.* 47 (1979).
8. Cloos, E. *Proc. natn. Acad. Sci. U.S.A.* **18,** 387–395 (1932).
9. Dennis, J. G. *Am. Ass. petrol. Geol. Mem.* **7,** 95–96 (1967).
10. Durney, D. W. & Ramsay, J. G. in *Gravity and Tectonics* (eds De Jong, K. A. & Scholten, R.) 67–96 (Wiley, New York, 1973).
11. Ramsay, J. G. *Nature* **284,** 135–139 (1980).
12. Grindley, G. W. & Browne, P. R. L. *Proc. 2nd U.N. Symp. on Development and Use of Geothermal Resources* 377–386 (New York, 1976).
13. Secor, D. T. *Am. J. Sci.* **263,** 633–646 (1965).
14. Heard, H. C. & Rubey, W. W. *Bull. geol. Soc. Am.* **77,** 741–760 (1966).
15. Sibson, R. H. *J. geol. Soc. Lond.* **133,** 191–213 (1977).
16. Hubbert, M. K. & Rubey, W. W. *Bull. geol. Soc. Am.* **70,** 115–166 (1959).
17. Hoek, E. & Bray, J. W. *Rock Slope Engineering* (Institute of Mining and Metallurgy, London, 1977).
18. Patterson, M. S. *Experimental Rock Deformation–The Brittle Field* (Springer, Berlin, 1978).
19. Anderson, E. M. *The Dynamics of Faulting* (Oliver & Boyd, Edinburgh, 1951).
20. Hubbert, M. K. & Willis, D. G. *Trans. Am. Inst. Min. Engr.* 210, 153–168 (1957).
21. Abdallah, A. *et al. Nature* **282** 17–23 (1979).
22. Narasimhan, T. N., Houston, W. N. & Nur, A. M. *Geology* **8,** 349–351 (1980).

Reworking in Upper Mesozoic and Cenozoic central Pacific deep-sea sediments

Jörn Thiede

Department of Geology, University of Oslo, PO Box 1047,
Blindern, Oslo 3, Norway

Data from deep-sea drilling cores from upper Mesozoic and Cenozoic central Pacific Ocean sediments have been used to estimate the frequency with which pelagic microfossils in sediments of different ages include species which have been displaced from older deposits. Intruding microfossils, which include planktonic foraminifera and calcareous nannofossils, and to a minor degree also radiolarians and diatoms, are most easily distinguished by not being contemporaneous with the sediments in which they are found. Presumably their occurrence arises from the mechanical reworking of oceanic sediments by bottom water currents. The data suggest that there were pulses of intensive reworking around 40 Myr BP, 30 Myr BP, 15 Myr BP, 5–10 Myr BP and during the Plio–Pleistocene; these pulses were interrupted by phases of only modest erosion on the Pacific Ocean floor. The absence of reworking of non-contemporaneous fossils in sediments older than 70 Myr suggests relatively sluggish deep-sea currents which were not able to erode deeply into the underlying Cretaceous and in certain areas Jurassic sediments.

Descriptions[1] of sediment columns probed by the Deep-Sea Drilling Project (DSDP) in the central Pacific quite frequently mention the occurrence of non-contemporaneous (reworked) microfossils which are older than the deposits in which they are included. The sites of occurrence (Fig. 1) are distributed over the entire western subtropical and tropical Pacific Ocean, whose basaltic basement is dominantly of Jurassic, Cretaceous and early Tertiary age[2], but there are also sites from which reworked material has not been reported, either because none exists or because it has been overlooked. The collected data[1] come from 35 DSDP sites drilled during 13 legs and are considered representative. Sediments with reworked fossils span a wide range of time (from late Jurassic–early Cretaceous to Quaternary) and depth (>1.5–<6 km), and most of them are located on the Mesozoic part of the Pacific Plate (Fig. 1). The poor coring record during the early DSDP phases and long hiatuses imply that deep-sea basin sites in the central Pacific Ocean are under-represented in comparison with elevated areas (aseismic rises and ridges). Given the wide area coverage of drill locations there is no bias towards a single deep-sea basin or towards one very special depositional environment. Many of the sites discussed here have been drilled on submarine elevations which in late Mesozoic times were at very shallow water depths, but which subsequently subsided several kilometres. The fact that the displaced fossils have come from pelagic sediments and that most of them have come from a depth interval above the calcite

0028-0836/81/080667—04$01.00

compensation depth (CCD) suggest intermediate (roughly equivalent to bathyal) source regions at the flanks of such submarine highs.

The distribution of displaced pelagic fossils in the central Pacific is based on 207 observations of the four major planktonic microfossil groups (diatoms, radiolarians, coccoliths, foraminifera) whose grain size spectra range from silt to sand-size material. About 80% of the observations are from occurrences of reworked calcareous nannofossils and planktonic foraminifera. The data are therefore biased towards the depositional environments of calcareous oozes above the CCD and under the influence of intermediate[3], but not the deepest, water masses of the oceanic water column.

The biostratigraphic zonations derived from pelagic microfossils are well correlated with absolute chronologies[4-6]. The difference between the age of reworked (zone) fossils and the surrounding sediment can therefore often be determined very precisely. The age difference between the reworked material and the time of displacement (age of the present host sediment) ranges in general between zero (the reworking of contemporaneous or pene-contemporaneous deposits) and a maximum value which is determined by the difference in age of the oldest eroded sediments and the time of deposition of the reworked components. The most surprising feature of the distribution of reworked pelagic fossils in the central Pacific (Fig. 2) is the almost complete lack of redeposition in the Mesozoic. The earliest occurrences are in Maastrichtian sediments, but there are few observations until the late middle Eocene when a major episode of erosion into older deposits began at the depth of the intermediate water masses of the central Pacific. The planktonic forminiferal faunas and calcareous nannofossil floras of many sites bear therefore ample evidence of widespread reworking of Eocene to Cretaceous faunas and floras[7]. The past 40 Myr have been a time of constant reworking of deep-sea sediments (Fig. 2) with peaks of erosive events $\sim$40 Myr BP, 30–32 Myr BP, 14–15 Myr BP, 7–8 Myr BP and during the past 5 Myr. The modest number of observations during 10–13 Myr BP and in particular 15–25 Myr BP and 33–38 Myr BP indicate a temporary decrease of intensity of erosive processes. A characteristic for the central Pacific (Fig. 2) is that only coeval sediments have been reworked during most of the Mesozoic (only Cretaceous).

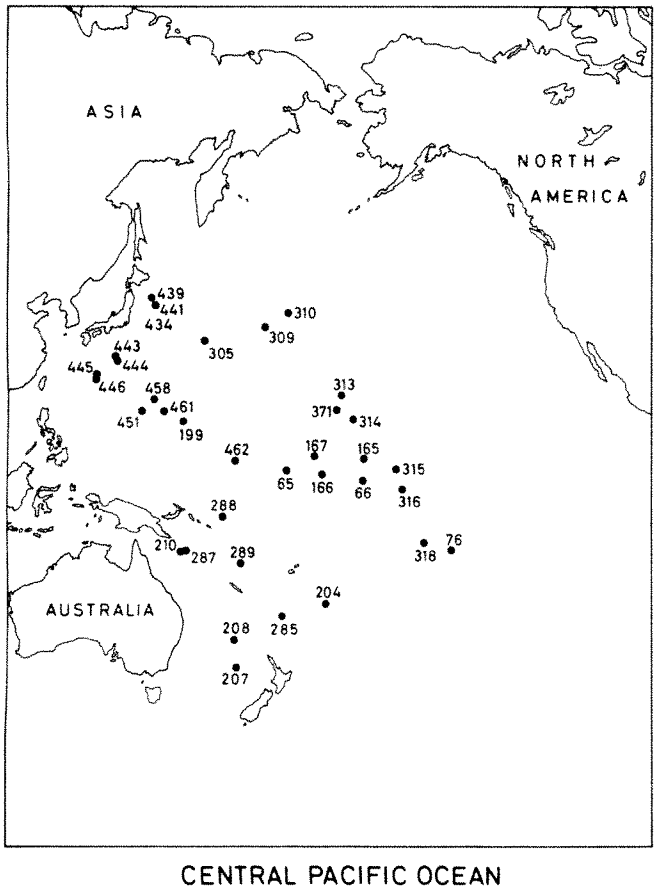


Fig. 1 Locations of deep-sea drilling sites with good coring records and reworked pelagic fossils in the tropical and subtropical central Pacific Ocean.

The time interval between the age of the eroded section and the time of final deposition of the reworked material can be large and the precision of this variable depends on the stratigraphic resolution[4-6]. The Mesozoic biostratigraphic zonations of deep-sea sediments are considerably less refined than Cenozoic ones[4]; thus Mesozoic and Cenozoic data from the central Pacific cannot be easily compared. However, the complete absence[1] of reworked non-contemporaneous pelagic microfossils in the pre-Maastrichtian parts of the central Pacific sediment section suggest that no older material has been subject to erosion in intermediate water depths during this time span. Age differences deviating considerably from zero only occur since Maastrichtian time (Fig. 2) when fossils from upper Campanian deposits were incorporated into chalks and calcareous oozes. The maximum age difference between the oldest reworked fossil components and the host sediment increased in a regular fashion from close to zero in Maastrichian to $\sim$38 Myr BP in late Eocene time when it reached 30–40 Myr. Between late Eocene and late Pliocene–early Pleistocene the maximum age difference fluctuated between 20 and 70 Myr. Maxima (Fig. 2) developed in the late Eocene, early Oligocene, early to middle Miocene and at the end of the Pliocene. Minima are confined to times when there are few observations of reworked pelagic material. The drastic decrease of the maximum difference in age of the reworked sediment components and the host sediment during early and late Pleistocene observed at Site 462 in the Nauru Basin[8] can also be seen in the data discussed here and seems to be a general feature of the redeposition of older pelagic deposits in the central Pacific.

Reworking and displacement of biogenic deep-sea sediment components imply mechanical erosion and transport of the fossil material[9] from an initial to the final site of deposition. Thus, the occurrence of reworked and displaced fossils in deep-sea sediments provides a simple tool for palaeoceanographic studies. Although reworked fossil material is abundant in certain horizons[10,11], it is generally only a minor fraction of the total sediment. Such fossils are therefore seldom described in detail in studies of the biostratigraphy and/or the depositional environment of Mesozoic and Cenozoic sedimentary sequences. The presence of reworked and displaced contemporaneous pelagic microfossils is often difficult to detect because of the apparently homogeneous nature of many deep-sea sediments and because displaced fossil assemblages cannot easily be discerned from autochthonous ones. The sedimentary structures in the modern ocean floor reflect intense bottom water current activity. The textural properties of deep-sea sediments[12] and sedimentary structures encountered in DSDP cores suggest that this also happened in the past, although it is difficult to evaluate the importance and the amount of reworked coeval material. The presence of displaced, non-contemporaneous pelagic sediment components is much easier to determine than that of contemporaneous material. They offer information about the nature, stratigraphy, palaeodepth of deposition and source of the eroded sediment section as well as about the current regimes of the depositional environments of the host sediment. Calcareous pelagic material has been eroded mostly from intermediate (deeper than neritic, but shallower than CCD) waters, purely siliceous displaced fossil assemblages from regions below the CCD.

The large difference between Mesozoic and Cenozoic redeposition (Fig. 2) of non-coeval reworked pelagic sediment components in the central Pacific is surprising. Maastrichtian events which reworked the directly underlying Campanian sediments represent the beginning of current regimes in the

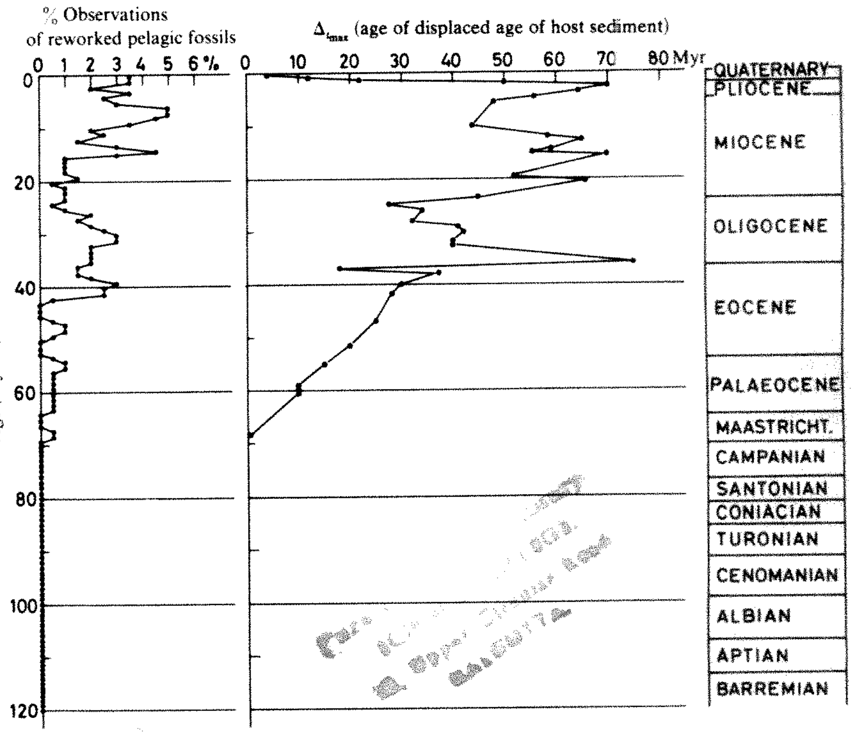


Fig. 2 Occurrence of non-contemporaneous reworked pelagic microfossils (calcareous nannofossils, planktonic foraminifers and radiolarians) in drill sites from the subtropical and tropical central Pacific Ocean (Fig. 1). Occurrence is expressed in per cent of total occurrences of reworked non-contemporaneous microfossils. $\Delta_{t_{\max}}$ is the maximum time difference per time interval between the age of the reworked and displaced pelagic fossils and of the host sediment.

intermediate water masses of an intensity which not only allowed transportation and displacement of coeval sediments, but also erosion of the underlying older strata. This is different from the South Atlantic palaeoenvironment where pelagic sediments as old as Coniacian have been displaced into deposits of Santonian age[13]. Although the available data are preliminary, the large quantitative differences of erosion and redeposition during pre-Maastrichtian and Maastrichtian–Pleistocene times indicate a fundamental change of the depositional palaeoenvironment of the deep central Pacific and probably of the world ocean in general. If the physical properties of deep-sea sediments[14] deposited during pre-Campanian time in the central Pacific were not drastically different from those laid down during the past 75 Myr, then their erosive potential should have been approximately the same. Moreover if diagenetic changes have affected the physical properties of the pre-Campanian then these diagenetic changes must have occurred soon after deposition of these sediments.

The lack of reworked pelagic microfossils older than their pre-Maastrichtian host sediments coincided with the repeated development of oxygen deficient depositional environments between 80 and 120 Myr BP. The very quiet and sluggish bottom water circulation which is suggested by the laminated sediments[15] of the anoxic lithologies supports the interpretation that a wide area of the pre-Campanian central Pacific in intermediate water depths has not experienced bottom water currents strong enough to erode older strata and to transport the resuspended sediment particles. Relatively rapidly flowing currents seem to have been confined to the upper few hundred metres of the water column, especially close to the equatorial divergence[15].

The number of occurrences of reworked non-coeval pelagic fossils (Fig. 2) remained modest from Maastrichtian to late middle Eocene times, but minor peaks during late Palaeocene and early middle Eocene represent short lived erosive events of a modest scale. The remainder of the Cenozoic was characterized by depositional palaeoenvironments where erosion in intermediate water depths was predominant. The temporal coincidence of important reworking with major changes of the oceanic bottom water temperature regimes (Fig. 3) as deduced from the δ^{18}O ratios of the benthic foraminifers[16,17], suggest that these episodes represent major changes in the bottom water current regimes due to the advection of cold polar probably Antarctic[8,19] to the intermediate waters in the central western Pacific Ocean. This advection resulted in more rapidly flowing currents which eroded the ocean floor and generated widespread hiatuses through reworking and mechanical redeposition of biogenic sediments rather than through dissolution. We do not know whether the pre-late middle Eocene reworking events can be explained by the formation of modest quantities of polar bottom waters much earlier than presumed hitherto which coincided with a significant temperature decrease in high latitudes in Maastrichtian times[20].

The time difference between age of reworked components and time of redeposition reflects the length of the stratigraphic column which was subject to erosion at the time of redeposition; it approached zero during Maastrichtian time, but increased fairly regularly throughout the Cenozoic to the early Pleistocene. The regular increase must mean that the entire sediment column which had been laid down since initiation of the deep

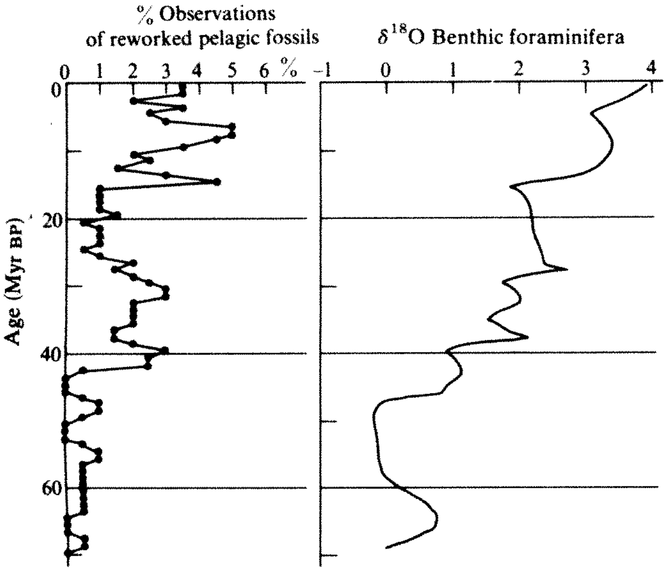


Fig. 3 Occurrence of non-contemporaneous reworked pelagic microfossils in the central Pacific Ocean and oxygen isotope ratios in benthic foraminifers through the past 70 Myr. Isotope data mostly from the North Pacific Ocean[17].

water current regime in the early Maastrichtian was subject to continuous erosion and redeposition. Deviations from the regular increase of the maximum time difference between the age of the reworked material and of the time of redeposition have been observed during early and late Oligocene as well as during late Miocene and Quaternary. Minima of age difference (Fig. 2) coincide with episodes of high pelagic accumulation rates[21,22], the maxima with the development of important hiatuses in the western Pacific Ocean[23]. The presence of reworked and displaced non-coeval fossils in deep-sea sediments implies mechanical transport and suggests that mechanical erosion rather than non-deposition or dissolution causes most of these hiatuses. The decrease of the maximum age difference between reworked and host sediments in the Central Pacific during the past 3–4 Myr suggests a reduction in the intensity of erosion in the intermediate part of the oceanic water column, despite the deteriorating climate. This is supported by accumulation rate data[24] from the polar South Pacific which indicate a decrease of erosion due to a decrease in current velocities over the past 3 Myr.

Changes of the global palaeoenvironment across the boundary from the late Mesozoic to the Cenozoic era caused: (1) an important decrease in the amount of hydrocarbons which formed on Earth[25]; (2) the initiation of a climatic regime which deteriorated from the equitable late Mesozoic mode to the Glacial mode towards the end of the Cenozoic; and (3) a catastrophy within the biosphere which led to the extinction of many forms of life on land and in the oceanic surface water masses. This development had previously been related to a gradual change of the global palaeoenvironment towards the end of the Cretaceous. More recently instantaneous events such as the sudden injection of brackish water masses from previously isolated Arctic Ocean[26] or the impact of a small celestial body[27] have been invoked to explain the sudden catastrophy at the Cretaceous–Tertiary boundary. However, the evidence from reworked pelagic fossils in the central Pacific Ocean indicates that an important component of the late Mesozoic palaeoenvironment, the current regimes of the intermediate and deep oceanic water masses, changed ~5 Myr earlier than the catastrophic events on land and in the pelagic surface waters of the oceans.

The data used in this report have been produced by the shipboard palaeontologists and sedimentologists of the DSDP Legs 7, 9, 17, 20, 21, 30, 32, 33, 57 to 61, their skilful work is acknowledged. We thank O. Eldholm (Oslo), and W. H. Berger for helpful comments. The Deep-Sea Drilling Project (DSDP) is operated by Scripps Institution of Oceanography with advice from the Joint Oceanographic Institutions.

Received 9 July; accepted 4 December 1980.

1. Thiede, J., Boersma, A., Schmidt, R. R. & Vincent, E. *Init. Rep. DSDP* Leg **62** (in the press).
2. Pitman, W. S., Larson, R. L. & Herron, E. M. *Magnetic Lineations of the Oceans, Map* (Geological Society of America, Boulder, 1974).
3. Reid, J. L. *The Johns Hopkins Oceanogr. Stud.* **2,** (1965).
4. van Hinte, J. E. *Am. Ass. petrol. Geol. Bull.* **60,** 498–516 (1976).
5. Hardenbol, J. & Berggren, W. A. *Am. Ass. petrol. Geol. Stud. Geol.* **6,** 213–234 (1978).
6. Berggren, W. A. & van Couvering, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **16,** 1–216 (1974).
7. Krasheninnikov, V. *Init. Rep. DSDP* Leg **62** (in the press).
8. Schlanger, S. O., Premoli-Silva, I., Rea, D. K. & Thiede, J. *Init. Rep. DSDP* Leg **61** (in the press).
9. Seibold, E. *Micropaleontology* **24,** 407–421 (1978).
10. Beckmann, J. P. *Init. Rep. DSDP* Leg **33,** 467–470 (1976).
11. Kaneps, A. G. *Init. Rep. DSDP* Leg **33,** 361–367 (1976).
12. van Andel, T. H. *J. Geol.* **81,** 434–457 (1974).
13. Premoli-Silva, I. & Boersma, A. *Init. Rep. DSDP* Leg **39,** 615–641 (1977).
14. Schlanger, S. O. & Douglas, R. G. *Int. Ass. Sediment. Spec. Publ.* **1,** 117–148 (1974).
15. Dean, W., Claypool, G. & Thiede, J. *Init. Rep. DSDP* Leg **62** (in the press).
16. Savin, S. M. *Ann. Rev. Earth planet. Sci.* **5,** 319–335 (1977).
17. Vincent, E. & Berger, E. H. in *The Sea* Vol. 7 (ed. Emiliani, C.) (Wiley-Interscience, New York, in the press).
18. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP* Leg **29,** 743–755 (1975).
19. Keigwin, L. D. *Nature* **287,** 722–725 (1980).
20. Hecht, A. D. *Foraminifera* **2,** 1–43 (1976).
21. van Andel, T. H., Heath, G. R. & Moore, T. C. *Geol. Soc. Am. Mem.* **143,** 1–84 (1975).
22. Worsley, T. R. & Davies, T. A. *Science* **203,** 455–456 (1979).
23. Moore, T. C., van Andel, T. H., Sancetta, C. & Pisias, N. *Micropaleontology* **24,** 113–138 (1978).
24. Ledbetter, M. T. & Huang, T. C. *Mar. Geol.* **36,** M21–M28 (1980).
25. Irving, E., North, F. K. & Couillard, R. *Can. J. Earth Sci.* **11,** 1–17 (1974).
26. Thierstein, H. R. & Berger, W. H. *Nature* **276,** 461–466 (1980).
27. Hsü, K. J. *Nature* **285,** 201–203 (1980).

Cenozoic palaeotemperature changes and planktonic foraminiferal speciation

Robert C. Thunell

Department of Geology and Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina, Columbia, South Carolina 29208

The interaction of plate movements and progressive climatic deterioration produced a constantly changing oceanic environment during the Cenozoic. The destruction of the Tethyan Seaway and the severing of low-latitude oceanographic connections during the past 65 Myr (ref. 1), coupled with the initiation of unrestricted circum-Antarctic circulation in the middle to late Oligocene[2], provided the necessary boundary conditions for the development of a global ocean marked by steep latitudinal temperature gradients. This is in contrast to the rather mild climate of the Mesozoic, which was characterized by an ocean with a homogeneous thermal structure and weak latitudinal temperature gradients[1]. Throughout the Cenozoic, marine plankton were constantly adapting to major changes in the oceanic environment. This can be seen, for example, in the continual development of latitudinal provincialization as a response to polar coolings[1,3]. Here I attempt to evaluate the relationship between palaeotemperature change and planktonic foraminiferal speciation during the past 65 Myr.

Earlier studies compared patterns of planktonic foraminiferal evolution with available palaeotemperature data[4–8]. Frerichs[7] suggested that rapid radiation and high diversity of planktonic foraminifera coincided with periods of warm climate, whereas Cifelli[5] and Lipps[6] both concluded that planktonic foraminiferal diversity was closely regulated by high-latitude marine climates. According to their models, when latitudinal temperature gradients are enhanced there is increased provincialization, while periods with globally uniform oceans are marked by low diversity and wide geographical distribution of species. Stehli *et al.*[9] and Cifelli[5] also proposed that the rate of evolution of new taxa is highest in regions of warmest surface waters. If true, marine sequences from tropical–subtropical regions should be an excellent record of the major features of planktonic foraminiferal evolution. Hart[10] suggested that not only is foraminiferal evolution closely related to palaeotemperature changes, but that the interative trends[5,11] observed during the Cenozoic "indicate successive attempts at colonizing deeper levels of the winter column".

The present data are based on a synthesis of Palaeocene to Recent planktonic foraminiferal distributions in marine and land-based records from the Caribbean region[12]. This compilation is based primarily on the results of Legs 4 (ref. 13) and 15 (refs 14, 15) of the Deep Sea Drilling Project, as well as information from the Antilles and northern South America[16–18]. The past 65 Myr has been divided into 43 planktonic foraminiferal zones[14,15], varying in duration between ~0.5 and 3.4 Myr (ref. 19). For each of these zones species diversity and the number of first and last appearances have been determined (Fig. 1). The species diversity data enable the total rate of change of planktonic foraminifera to be estimated in terms of number of species per Myr (Fig. 1).

The $\delta^{18}O$ record used in this study is a compilation of previously published benthic foraminiferal data[20]. The Cenozoic $\delta^{18}O$ record reveals that there has been a persistent cooling of the oceans since the late Palaeocene. Superimposed on this trend are a series of distinct climatic and oceanographic events, including: (1) the development of the psychrosphere near the Eocene–Oligocene boundary[21]; (2) the establishment of an unrestricted circum-Antarctic current in the middle to late Oligocene[2]; (3) the development of a permanent Antarctic ice cap in the middle Miocene[22]; and (4) the onset of Northern Hemisphere glaciation in the late Pliocene[23–26].

0028–0836/81/080670—03$01.00

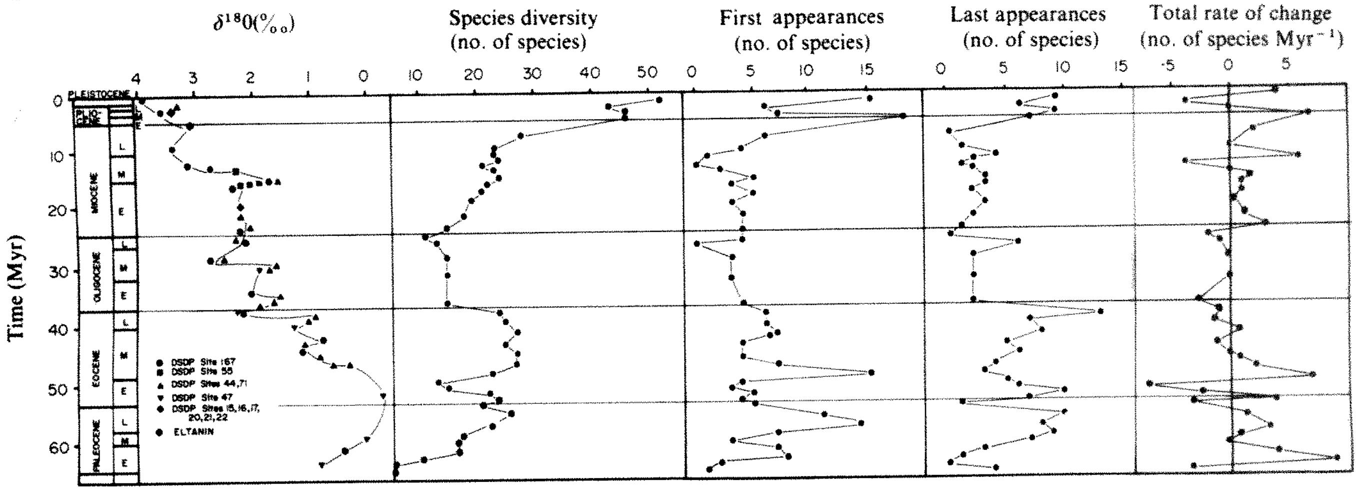


Fig. 1 Comparison of Cenozoic $\delta^{18}O$ record with various faunal parameters reflecting the diversity of planktonic foraminifera. The $\delta^{18}O$ curve is from ref. 20 and is a composite record based on analyses of benthic foraminifera from different deep-sea sedimentary sections. The faunal data base is from ref. 12. Geological epochs and an absolute time scale after Berggren[19] are plotted on the left.

The palaeotemperature record indicates that a gradual warming occurred throughout the Palaeocene. Associated with this warming is an increase in species diversity (from <5 species at the base of the Palaeocene to >25 species in the late Palaeocene), with the morphologically simple globigerine forms that survived the massive extinctions at the end of the Mesozoic giving rise to the more advanced globorotaliid forms[5,27]. There are two distinct episodes of rapid origination within the Palaeocene, one in the early and one in the late Palaeocene. These are reflected in the total rate of change of the fauna, with the Palaeocene generally being a time of net gain in overall number of species.

Following the diversity high in the late Palaeocene there was a rapid decrease in species number with a low being reached by the end of the early Eocene. There is a sharp decline in the number of first appearances during this time interval, while the number of last appearances remains relatively high. The result is a net loss in species with the latter part of the early Eocene being marked by the greatest decrease for the entire Cenozoic. This is when the long-term cooling trend for the Cenozoic probably began, although there is no abrupt change in the isotopic record with which we can specifically correlate the observed faunal events.

A second major increase in species diversity occurred in the Palaeogene, extending from the middle Eocene through the end of the Eocene, where first appearances exceed extinctions. The major evolutionary feature during this period was the introduction of hastigerine, orbuline and truncorotaliid forms[5,27]. The palaeotemperature record indicates a brief warming during the middle Eocene followed by relatively stable climatic conditions throughout the remainder of the Eocene. (A similar pattern is shown by isotopic work reported in refs 22, 28, 29.)

One of the most dramatic events in the Tertiary oxygen isotopic record is associated with the Eocene–Oligocene boundary (~38 Myr ago). This rapid isotopic change has been attributed to a cooling of Antarctic waters and the development of cold bottom water circulation (psychrosphere) similar to that of the present day[21]. The response of planktonic foraminifera to this new oceanographic regime is manifested as a sudden drop in species diversity, (from ~25 to 15 species) with the largest number of last appearances for the entire Cenozoic. By the end of the Eocene, all globorotaliid, hastigerine, orbuline and truncorotaliid forms had become extinct. As at the Cretaceous–Tertiary boundary, the simple globigerine forms were among the few survivors of the Eocene–Oligocene event. A similar crisis in benthic foraminifera was reported by Kennett and Shackleton[21], although Corliss[30] suggests that the benthic turnover may have occurred in more gradual steps.

Following the climatic deterioration at the Eocene–Oligocene boundary, fairly stable oceanographic conditions were maintained until the middle Oligocene. This stability is reflected in a planktonic foraminiferal fauna with uniformly low diversity. The simple globigerine assemblage present at this time has a global geographical distribution indicating the absence of any strong latitudinal thermal barriers.

Another abrupt isotopic shift occurs in the middle Oligocene. This isotopic event coincides with the initiation of unrestricted circum-Antarctic flow[2], and may reflect a high-latitude cooling. This isotopic event coincided with the largest eustatic drop in sea level of the entire Cenozoic[31]. Each faunal parameter suggests that there was a delay in the response of planktonic formaminifera to this oceanographic event. Shortly after the isotopic shift, planktonic foraminiferal species diversity reached a low point, the number of first appearances dropped to zero, and there was an increase in last appearances. This offset in the isotopic and faunal records may reflect a lag in the response of planktonic foraminifera to changing surface conditions, or it may be a stratigraphic artefact. The ubiquitous globigerine forms continue to dominate the planktonic foraminiferal fauna of the middle and late Oligocene.

This diversity low in the late Oligocene is followed by a very uniform increase in the number of planktonic foraminiferal species that begins in the early Miocene and continues throughout the remainder of the Neogene. Species diversity increases from ~10 species in the late Oligocene to 25 species by the end of the early Miocene. Cifelli[5] considered this to be one of two major expansions in planktonic foraminifera that occurred during the Cenozoic (the other following the massive extinctions at the Cretaceous–Tertiary boundary), with the modern fauna beginning to evolve at this time. Phylogentically different lineages of globigerinoids, globorotaliids, hastigerines and orbulines reappeared during the early Miocene. The palaeotemperature record indicates a slight warming at the end of the Oligocene, followed by stable conditions during the early Miocene. Associated with these uniformly warm conditions of the early Miocene were a constant rate of first appearances and a low level of last appearances resulting in net gain in total fauna.

The relatively stable oceanographic conditions of the early Miocene ended abruptly in the middle Miocene (about 14–12 Myr ago) with the development of a permanent Antarctic ice cap[22]. This major climatic event is manifested by the largest isotopic shift in the Cenozoic record, and further enhanced latitudinal provincialization of all planktonic microfossil groups[3]. First appearances drop to zero, while the number of extinctions increases slightly in the middle Miocene producing a brief but relatively large net loss in species. Throughout this period species diversity remains fairly stable, fluctuating between 21 and 24 species.

The most rapid increase in species diversity during the past 65 Myr occurred in the early Pliocene. Within several million years, the number of planktonic foraminiferal species increased from <30 in the late Miocene to >45 in the early Pliocene. Since the middle Miocene, the world had been in a glacial state but the isotope record indicates that a slight retreat of the Antarctic ice cap may have occurred in the early Pliocene. The rapid increase of planktonic foraminifera at this time may have been a response to this brief amelioration.

The most recent major climatic threshold was the initiation of Northern Hemisphere glaciation in the late Pliocene (~3.0 Myr ago)[23-26]. This probably intensified existing latitudinal temperature gradients. This climatic change immediately caused a dramatic drop in number of first appearances, a slight decline in species diversity and a short-term net loss in species.

Since the inception of the northern ice sheets, there have been periodic oscillations between glacial and interglacial climates. However, these rapidly changing climatic conditions have had very little adverse effect on planktonic foraminifera. The general pattern has been for the geographic ranges of individual species to expand and contract in response to changes in latitudinal temperature gradients, without any appreciable loss of species. Once planktonic foraminifera became adapted to the effects of a Northern Hemisphere ice sheet, species diversity increased and reached a maximum for the entire Cenozoic (>50 species).

This comparison of the Cenozoic palaeotemperature record and species diversity in planktonic foraminifera confirms the previously reported[4,7,8] close relationship between palaeoclimatic change and planktonic foraminiferal evolution. During the past 65 Myr, there have been five periods during which there was a net loss of the total planktonic foraminiferal population. These periods of net loss correspond closely to the major climatic events of the Cenozoic, and include the initial phase of climatic deterioration in the early Eocene, the development of the psycosphere near the Eocene–Oligocene boundary, the initiation of unrestricted circum-Antarctic flow in the middle to late Oligocene, the build-up of a permanent Antarctic ice cap in the middle Miocene and the onset of Northern Hemisphere glaciation in the late Pliocene. Following each of these major climatic events there was generally an expansion of planktonic forminifera associated with milder climate. The species diversity record clearly shows that there were two major radiations during the Cenozoic, one beginning in the early Palaeocene and one beginning in the early Miocene[5].

I thank Drs W. A. Berggren, B. H. Corliss and J. P. Kennett for helpful comments. This work was partially supported by NSF grant OCE79-23736 and a Research and Productive Scholarship Grant from the University of South Carolina.

Received 15 September; accepted 4 December 1980.

1. Berggren, W. A. & Hollister, C. D. *Soc. Econ. Paleont. Mineral., Spec Publ.* **20,** 126–186 (1974); *Tectonophysics* **38,** 11–48 (1977).
2. Kennett, J. P. *J. geophys. Res.* **82,** 3843–3860 (1977).
3. Kennett, J. P. *Mar. Micropaleont.* **3,** 301–345 (1978).
4. Berggren, W. A. *Micropaleontology* **15,** 351–365 (1969).
5. Cifelli, R. S *Syst. Zool.* **18,** 154–168 (1969).
6. Lipps, J. H. *Evolution* **24,** 1–22 (1970).
7. Frerichs, W. E. *J. Paleont.* **45,** 963–968 (1971).
8. Fischer, A. G. & Arthur, M. A. *Soc. Econ. Paleont. Mineral. Spec. Publ.* **25,** 19–50 (1977).
9. Stehli, F. G., Douglas, R. G. & Newell, N. D. *Science* **164,** 947–949 (1969).
10. Hart, M. B. *Nature* **286,** 252–254 (1980).
11. Steineck, P. L. & Fleischer, R. L. *J. Paleont.* **52,** 618–635 (1978).
12. Saunders, J. B. *et al. Init. Rep. DSDP* Leg **15,** 796–771 (1973).
13. Bolli, H. M. *Init. Rep. DSDP* Leg **4,** 557–643 (1970).
14. Bolli, H. M. & Premoli Silva, I. *Init. Rep. DSDP* Leg **15,** 475–497 (1973).
15. Premoli Silva, I. & Bolli H. M. *Init. Rep. DSDP* Leg **15,** 499–547 (1973).
16. Blow, H. W. *Int. Conf. Plank. Microfossils*, Geneva, 199–422 (1969).
17. Bolli H. M. *Bull. U.S. Nat. Mus.* **215,** 97–123 (1957); **215,** 155–172 (1957).
18. Bolli, H. M. & Bermudez, P. J. *Bull. Ass. Venez. geol. Miner. Petrol.* **8,** 119–149 (1965).
19. Berggren, W. A. *Lethaia* **5,** 195–215 (1972).
20. Savin, S. M. *A. Rev. Earth planet. Sci.* **5,** 319–355 (1977).
21. Kennett, J. P. & Shackleton, N. J. *Nature* **260,** 513–515 (1976).
22. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP* Leg **29,** 743–755 (1975).
23. Berggren, W. A. *Init. Rep. DSDP* Leg **12,** 953–963 (1972).
24. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP* Leg **29,** 801–807 (1975).
25. Shackleton, N. J. & Opdyke, N. D. *Nature* **270,** 216–217 (1977).
26. Keigwin, L. D. & Thunell, R. C. *Nature* **282,** 294–296 (1979).
27. Berggren, W. A. *Proc. 11th Plank. Conf.*, Rome, 41–56 (1971).
28. Boersma, A. & Shackleton, N. J. *Init. Rep. DSDP* Leg, **39,** 911–924 (1977).
29. Vergnaud-Grazzini, C. *Init. Rep. DSDP* Leg **47,** 507–511 (1979).
30. Corliss, B. H. *Nature* **282,** 63–65 (1979).
31. Vail, P. R., Mitchum, R. M. Jr & Thompson, S. *Mem. Am. Ass. petrol. Geol.* **26,** 83–97 (1977).

Rates of emission of H_2S from plants and patterns of stable sulphur isotope fractionation

W. E. Winner*, C. L. Smith*, G. W. Koch*, H. A. Mooney*, J. D. Bewley† & H. R. Krouse‡

* Department of Biological Sciences, Stanford University, Stanford, California 94305
† Department of Biology and ‡Department of Physics, University of Calgary, Calgary, Alberta, Canada T2N 1N4

In all green plants both HSO_3^- and SO_4^{2-} are reduced during photosynthesis, and many species can emit the product as H_2S gas. Plants will emit H_2S after fumigation with acute doses of SO_2 gas[1,2], irrigation with 5% K_2SO_4 solutions[3], or immersion of roots in HSO_3^- and SO_4^{2-} solutions[4]. If such emissions of H_2S by plants occur in nature, then this process adds sulphur to the atmosphere and contributes to the global sulphur budget. Oxidation of biogenic H_2S in the atmosphere may also contribute to the formation of acid rain. We report here that irradiated plants with roots immersed in HSO_3^- and SO_4^{2-} solutions at concentrations found in nature emit H_2S from their leaves, and that the fractionation of stable isotopes of sulphur during H_2S emission may be useful for identifying atmospheric sulphur that has been generated by photosynthetic reduction of sulphur.

Rates of emission of H_2S were measured for 10-week-old soybean plants (*Glycine max* L. Merr. cvs. Kent, Peking and York) raised in growth cabinets. Days were 12 h long, photosynthetically active radiation (PAR) was about 600 μE m^{-2} s^{-1}, humidity was 50–60% and the temperature was 15 °C at night and 20 °C during the day. Plants were prepared for experiments by removing them from potting media, washing and trimming their roots, and placing them in a flask with distilled water. The flask and plants were placed in a flow-through Plexiglass chamber, the upper portion of which was isolated by a diaphragm. Free space around the canopy was reduced by foam spacers inserted into the chamber, and air around the leaves was mixed by a propeller, driven by an electrical motor mounted outside the chamber. The roots were aerated by one airstream, and another airstream was passed through activated charcoal and used to flush the air space around the leaves. The chamber wall, diaphragm, foam spacers and propeller were lined with FEP Teflon to reduce H_2S absorption. After 30 min of incubation, the chamber was sealed, and air leaving it was analysed for H_2S with an Interscan Model 1170 sp H_2S analyser. Sampling was limited to 15 min to prevent water vapour from condensing in the chamber. Radiation by Sylvania Cool Lux bulbs was filtered through water and was 500 μE m^{-2} s^{-1} (300–760 nM) inside the chamber.

Plants in distilled water were tested for H_2S emission, then Na_2SO_4 (Baker, analytical grade) was added to the flask to increase the amounts of SO_4^{2-} available to the roots. Concentrations of SO_4^{2-} in the test solutions were 0.04, 0.1 and 1.0 mM. The highest concentration used is thought to be optimal for growth of agricultural crops[5], and the lowest concentration is similar to that of acid rain in the eastern US[6]. After treatment with the series of SO_4^{2-} solutions, plants were dried in an oven and rates of H_2S emission calculated on a plant dry-weight basis using the formula:

$$J_{H_2S} = XFW^{-1} \qquad (1)$$

where J_{H_2S} is the H_2S emission rate (ng g^{-1} s^{-1}), X the H_2S concentration leaving the chamber (ng cm^{-3}), F the airstream flow rate (cm^3 s^{-1}) and W plant dry weight (g).

0028-0836/81/080672—02$01.00

A similar system was used to determine the fractionation of stable sulphur isotopes during emission. We used 4-week-old cucumber plants (*Cucumis sativus* L. cv. Straight Eight) grown in a greenhouse, because these were found to emit H_2S in previous experiments[4]. The isotopic composition of the test solutions (0.1–0.2 M HSO_3^- or SO_4^{2-}) was known. To obtain sufficient H_2S for analysis with a mass spectrometer, sulphur ions were at a higher concentration in test solutions for these experiments than for determining rates of emission. The H_2S emitted by the plants was trapped as Ag_2S by filtering the air leaving the chamber through $AgNO_3$ deposited on quartz wool. This procedure was repeated several times with fresh plants and fresh test solution until several mg of Ag_2S were collected. The H_2S traps showed no selectivity for sulphur isotopes. After these experiments, leaves were oven dried, and sulphur was combusted to SO_4^{2-} in a Parr bomb. Washings were treated with Ba^{2+} to precipitate $BaSO_4$ and then chemically reduced to yield Ag_2S[7]. $\delta^{34}S$ values (Table 1) for Ag_2S from both leaf sulphur and H_2S gas were determined by isotope ratio mass spectrometry by reacting Ag_2S with Cu_2O at 900 °C, to form SO_2.

H_2S was detected from plants in all test solutions containing SO_4^{2-} (Fig. 1). Even plants in the most dilute solutions (0.04 mM) had a mean emission rate of 0.036 ng g^{-1} s^{-1}. Rates of emission increased as the concentration of SO_4^{2-} in the test solution increased, and the highest rate recorded was 0.07 ng g^{-1} s^{-1} for plants treated with a test solution of 1.0 mM SO_4^{2-}. H_2S emission rates persisted for several hours when experiments were prolonged. Emission rates observed here are lower than the 10 μg g^{-1} s^{-1} observed for detached cucumber leaves treated with 25–100 mM SO_4^{2-} and higher than the 0.6 pg g^{-1} s^{-1} observed for spruce seedlings[3]. When the lights were turned off after two experiments, emission rates declined within 15 min to about 30% of the highest rate. Other workers have also observed the light dependence of SO_4^{2-} reduction[8,9] and H_2S emission[1,2,4]. No H_2S was emitted from plants placed in distilled water, and there seemed to be no consistent differences in emission rate between the three cultivars at any single concentration of SO_4^{2-}.

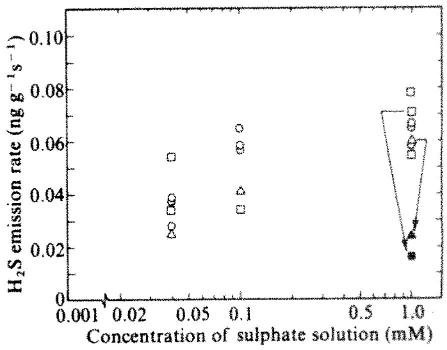


Fig. 1 H_2S emission rates (plant dry-weight basis) from irradiated soybean plants (open symbols) and from two plants transferred to darkness for a period of 15 min (closed symbols). Emission rates were monitored for three cultivars: ○, Kent; □, York; △, Peking.

If SO_4^{2-} is available to plants at a concentration of 0.05 mM and they emit H_2S at a rate of 0.04 ng g^{-1} s^{-1}, then the amount of H_2S produced worldwide from photosynthetic reduction of SO_4^{2-} could be 5.4×10^{13} g yr^{-1}. This constitutes about 50% of the amount needed to balance the global sulphur budget[10,11]. Our calculations were made by estimating global biomass at 1.85×10^{12} dry tons[12], by assuming that about 5% of this weight is composed of leaves[13], and then by averaging the daily photoperiod to 12 h. Plants may emit H_2S at greater rates if SO_4^{2-} concentrations are generally greater than 0.05 mM. In addition, H_2S emitted by plants may be oxidized in the atmosphere and eventually returned to soils in the form of acid rain. A single sulphur atom could conceivably be cycled many times between the soil and atmosphere by the processes of H_2S emission from plants and acid rain.

H_2S emitted by plants contained more ^{32}S and less ^{34}S than either the leaves or the test solutions (Table 1). δ^{34} S values differed by −9.7‰ from the SO_4^{2-} source and ~ −16.8‰ from the HSO_3^- source. The more extreme isotope fractionation pattern observed for HSO_3^- was also observed when both oxidized sulphur ions were supplied to plants in one solution. These results suggest that when both HSO_3^- and SO_4^{2-} are available, the emitted H_2S originates from the HSO_3^- source.

Photosynthetic sulphur-reducing bacteria show similar ^{32}S selectivity during H_2S formation[14]. Thus, depletion of ^{34}S resulting from isotope fractionation during H_2S formation may be common to chloroplasts and some prokaryotes. The ratio of isotopic rate constants for S–O bond rupture during inorganic reduction is similar ($k_{32}/k_{34} = 1.022$)[15] to the fractionations seen in our experiments.

We suggest that stable sulphur isotope analysis may indicate whether or not atmospheric sulphur compounds have originated from vegetation. Atmospherics derived from H_2S emissions from plants should be enriched in ^{32}S compared with stable sulphur isotopes that are available to plant roots. These fractionation patterns are independent of the oxidation state of sulphur in the air. We have observed that the $\delta^{34}S$ values of plants reflect $\delta^{34}S$ values of sulphur in their environment[16–18], and we suggest that plants that emit H_2S should become enriched in ^{34}S relative to their sulphur source.

This research was supported by NRC (Canada) grants to J.D.B. and H.R.K. and a grant to H.A.M. from the Electrical Power Research Institute (USA). We thank Dr Peter Vitousek for commenting on an earlier draft of this manuscript.

Table 1 Sulphur isotope abundances in test solutions, cucumber leaves and emitted H_2S on exposure of plant roots to HSO_3^- and SO_4^{2-} solutions of different concentrations

	$\delta^{34}S$		
	Solution*	Leaves	H_2S
Control	—	+10.7	—
Treatment with:			
0.1 M HSO_3^-	+2.2	+4.5	−14.5
0.2 M HSO_3^-	+2.2	+2.7	−14.6
0.2 M SO_4^{2-}	+4.0	+3.1	−5.7
0.1 M HSO_3^- + 0.1 M SO_4^{2-}	+3.1	+3.9	−13.4

Using accepted practice, $\delta^{34}S$ values are defined with respect to meteoritic troilite as

$$\delta^{34}S = [^{34}S/^{32}S]_{sample}/[^{34}S/^{32}S]_{meteorite} - 1 \times 10^3$$

Control; no H_2S was emitted from plants treated with distilled water.
* $\delta^{34}S$ value before experimentation with plants.

Received 29 September; accepted 3 December 1980.

1. Decormis, L. *C. r. hebd Séanc Acad. Sci.* D **266**, 683–685 (1968).
2. DeCormis, L. *Proc. 1st Eur. Congr. on the Influence of Air Pollution on Plants and Animals, Wageningen*, 75–78 (1969).
3. Spaleny, J. *Pl. Soil* **48**, 557–563 (1977).
4. Wilson, L. G., Bressan, R. A. & Filner, P. *Pl. Physiol.* **61**, 184–189 (1978).
5. Anderson, J. W. *Sulphur in Biology* Studies in Biology no. 101. (Arnold, London, 1978).
6. Likens, G. E. & Bormann, F. H. *Science* **184**, 1176–1179 (1974).
7. Thode, H. G., Monster, J. & Dunford, H. B. *Geochim. cosmochim. Acta* **25**, 159–174 (1961).
8. Schmidt, A. & Trebst, A. *Biochim. biophys. Acta* **180**, 529–535 (1969).
9. Trebst, A. & Schmidt, A. *Prog. Photosyn. Res.* **3**, 1510–1516 (1969).
10. Kellogg, W. W., Cadle, R. D., Allen, E. R., Lazrus, A. L. & Martell, E. A. *Science* **175**, 587–596 (1972).
11. Nguyen, B. D., Gaudry, A., Bonsang, B. & Lambert, G. *Nature* **275**, 637–639 (1978).
12. Whittaker, R. H. *Communities and Ecosystems*, 83 (Macmillan, New York, 1970).
13. Odum, E. P. *Fundamentals of Ecology*, 376 (Saunders, Philadelphia, 1971).
14. Kemp, A. L. W. & Thode, H. G. *Geochim. cosmochim. Acta* **32**, 71–91 (1958).
15. Harrison, A. G. & Thode, H. G. *Trans. Faraday Soc.* **53**, 1648–1651 (1957).
16. Krouse, H. R. *Nature* **265**, 45–46 (1977).
17. Winner, W. E., Bewley, J. D., Brown, H. M. & Krouse, H. R. *Oecologia* **36**, 351–361 (1978).
18. Case, J. W. & Krouse, H. R. *Oecologia* **44**, 248–257 (1980).

The stability of real ecosystems

P. Yodzis

Department of Zoology, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Most current theoretical work on whole ecological communities is based on a conception of community dynamics (see, for example, refs 1, 2) in which the community resides in a neighbourhood of equilibrium. However, there is very little evidence that this is a realistic viewpoint. To test this fundamental assumption 'empirically', I have constructed here plausible community matrices[1,2] corresponding to 40 real food webs[3] and studied their local stability properties. I find that stability (in the sense of tending to return to the presumed equilibrium after a small perturbation) is far more likely if the interaction strengths are chosen in accord with the nature of the particular organisms in each food web, rather than strictly at random. Either this is a monumental coincidence, or the equilibrium viewpoint really is appropriate for quite a few real communities. Additional results of this analysis are that stability generally requires a degree of intraspecific interference on the part of some consumer species, whereas interspecific interference tends to exert a destabilizing influence.

As emphasized by Pimm and Lawton[4], element A_{ij} of the community matrix gives the *per capita* effect of species j on species i in a neighbourhood of the presumed equilibrium. This is an important observation, for it enables one to make plausible guesses about the relative strengths of some interactions. Thus, the *per capita* effect of an insect population on a tree population is much smaller than the *per capita* effect of trees on insects, and so on[4]. An additional advantage of this approach is that it treats the community matrix as a truly local object, which in reality it is. Thus, the results reported here are completely independent of any assumptions about global dynamics; in particular, these are not Lotka–Volterra models.

For any real food web, we can construct a plausible community matrix by specifying four real, positive parameters f_1, s_1, f_2, s_2, with f_1 and f_2 chosen from the interval $[0, 1]$, and applying the following rules (which also serve to define the parameters f_1, s_1, f_2, s_2): (1) Trophic interactions. These matrix elements are chosen as by Pimm and Lawton[4] when their discussion applies. Cases not covered by Pimm and Lawton are treated in the spirit of their approach. It is in this step that a pattern of interaction strengths is arrived at which expresses the nature of the particular organisms in the food web. (2) Intraspecific interference. If species i is a primary producer, matrix element A_{ii} is chosen at random from the interval $(-s_1, 0)$. In addition, a fraction f_1 of the consumer species, chosen at random, is given diagonal matrix elements chosen at random from the interval $(-s_1, 0)$. The diagonal matrix elements for all other consumer species are set equal to zero. (3) Interspecific interference. If two consumers i and j have a prey species in common, then these species may indulge in interspecific interference, and I will call the ordered pair (i, j) 'potential competitors'. (Interspecific interference can also arise from other causes, such as competition for nest sites, but I will neglect such interactions here.) For a fraction f_2 of the potential competitors, chosen at random, the corresponding matrix elements A_{ij} are decreased from the values assigned in step (1) by amounts chosen at random from the interval $(0, s_2)$.

The food webs used in this study are the 40 webs recently compiled from the published literature by Briand[3], which is the most thorough existing community food web collection. For a variety of parameter sets $\{f_1, s_1, f_2, s_2\}$, 100 plausible community matrices were constructed for each Briand food web, using the above rules and a pseudorandom number generator, and the fraction of stable community matrices corresponding to each food web was computed. The results of some of these computations are displayed in Tables 1–4.

Table 1 shows the effect of increasing the strength s_1 of the intraspecific interference (relative to, of course, the strength of trophic interactions). In these computations, $f_1 = 1$ (all species in the community experience intraspecific interference) and $f_2 = 0$ (there is no interspecific interference). It is readily seen that stronger intraspecific interference favours stability, just as one expects.

Table 2 shows the effect of increasing the fraction f_1 of consumer species which engage in intraspecific interference. In these computations $f_2 = 0$ again, and $s_1 = 10^6$. Because of this large value for s_1 and in view of Tables 1 and 3, the values given in Table 2 can be viewed as upper bounds on the fraction of stable systems. It appears that at least one-tenth and perhaps as much as one-half of the consumer species in these communities need to exhibit some degree of intraspecific interference if the equilibrium viewpoint is to apply.

Pimm and Lawton[4] use the term 'self-limitation' for what I am calling intraspecific interference (namely, negative A_{ii}), and they

Table 1 Effect of changes in the strength of intraspecific competition on stability

Case no.	Community	Fraction of community matrices which are stable			
		$s_1 = 0.1$	$s_1 = 1$	$s_1 = 10$	$s_1 = 100$
1	Cochin estuary	0.00	0.12	0.99	1.00
2	Krysna estuary	0.00	0.64	1.00	1.00
3	Long Island estuary	0.01	0.49	1.00	1.00
4	California salt marsh	0.00	0.52	1.00	1.00
5	Georgia salt marsh	1.00	1.00	1.00	1.00
6	California tidal flat	0.00	0.07	0.99	1.00
7	Narragansett Bay	0.00	0.00	0.93	1.00
8	Bissel Cove marsh	0.00	0.07	0.96	1.00
9	Lough Ine rapids	0.01	0.43	0.99	1.00
10	Exposed intertidal (New England)	0.41	0.89	1.00	1.00
11	Protected intertidal (New England)	0.13	0.59	1.00	1.00
12	Exposed intertidal (Washington State)	0.00	0.41	0.99	1.00
13	Protected intertidal (Washington State)	0.00	0.19	1.00	1.00
14	Mangrove swamp (Station 1)	0.02	0.20	0.98	1.00
15	Mangrove swamp (Station 2)	0.01	0.24	0.99	1.00
16	Pamlico River	0.01	0.45	1.00	1.00
17	Marshallese reefs	0.00	0.39	0.99	1.00
18	Kapingamarangi atoll	0.00	0.33	1.00	1.00
19	Moosehead Lake	0.01	0.74	1.00	1.00
20	Antarctic pack ice zone	0.00	0.09	0.95	1.00
21	Ross Sea	0.00	0.12	0.95	1.00
22	Bear Island	0.00	0.01	0.92	1.00
23	Canadian prairie	0.00	0.19	0.97	1.00
24	Canadian willow forest	0.07	0.92	1.00	1.00
25	Canadian aspen forest	0.12	0.95	1.00	1.00
26	Aspen parkland	0.00	0.28	0.98	1.00
27	Wytham wood	0.00	0.12	0.97	1.00
28	New Zealand salt-meadow	0.14	0.78	1.00	1.00
29	Arctic seas	0.00	0.07	0.97	1.00
30	Antarctic seas	0.00	0.47	1.00	1.00
31	Black Sea epiplankton	0.00	0.00	0.82	1.00
32	Black Sea bathyplankton	0.00	0.00	0.76	1.00
33	Crocodile Creek	0.01	0.87	0.99	1.00
34	River Clydach	0.00	0.07	0.97	1.00
35	Morgan's Creek	0.00	0.71	1.00	1.00
36	Mangrove swamp (Station 6)	0.00	0.26	1.00	1.00
37	California sublittoral	0.01	0.76	1.00	1.00
38	Lake Nyasa rocky shore	0.00	0.34	1.00	1.00
39	Lake Nyasa sandy shore	0.00	0.63	1.00	1.00
40	Malaysian rain forest	0.00	0.67	1.00	1.00

0028-0836/81/080674—03$01.00

Table 2 Effect of changes in the frequency of intraspecific competition on stability

Case no.	Community	Fraction of community matrices which are stable			
		$f_1=0.0$	$f_1=0.1$	$f_1=0.5$	$f_1=1.0$
1	Cochin estuary	0.00	0.00	0.51	1.00
2	Krysna estuary	0.00	0.05	0.60	1.00
3	Long Island estuary	0.00	0.03	0.56	1.00
4	California salt marsh	0.00	0.04	0.62	1.00
5	Georgia salt marsh	1.00	1.00	1.00	1.00
6	California tidal flat	0.00	0.01	0.42	1.00
7	Narragansett Bay	0.00	0.00	0.32	1.00
8	Bissel Cove marsh	0.00	0.02	0.53	1.00
9	Lough Ine rapids	0.00	0.01	0.44	1.00
10	Exposed intertidal (New England)	0.53	0.61	0.89	1.00
11	Protected intertidal (New England)	0.15	0.16	0.72	1.00
12	Exposed intertidal (Washington State)	0.00	0.00	0.27	1.00
13	Protected intertidal (Washington State)	0.00	0.00	0.29	1.00
14	Mangrove swamp (Station 1)	0.11	0.34	0.85	1.00
15	Mangrove swamp (Station 2)	0.02	0.16	0.61	1.00
16	Pamlico River	0.21	0.32	0.79	1.00
17	Marshallese reefs	0.00	0.00	0.38	1.00
18	Kapingamarangi atoll	0.00	0.02	0.51	1.00
19	Moosehead Lake	0.00	0.01	0.41	1.00
20	Antarctic pack ice zone	0.00	0.01	0.44	1.00
21	Ross Sea	0.00	0.00	0.38	1.00
22	Bear Island	0.00	0.00	0.21	1.00
23	Canadian prairie	0.00	0.02	0.44	1.00
24	Canadian willow forest	0.00	0.10	0.65	1.00
25	Canadian aspen forest	0.00	0.00	0.11	1.00
26	Aspen parkland	0.00	0.00	0.21	1.00
27	Wytham wood	0.00	0.02	0.50	1.00
28	New Zealand salt-meadow	0.00	0.00	0.05	1.00
29	Arctic seas	0.00	0.00	0.13	1.00
30	Antarctic seas	0.00	0.00	0.23	1.00
31	Black Sea epiplankton	0.00	0.00	0.12	1.00
32	Black Sea bathyplankton	0.00	0.00	0.08	1.00
33	Crocodile Creek	0.00	0.00	0.34	1.00
34	River Clydach	0.00	0.00	0.55	1.00
35	Morgan's Creek	0.00	0.00	0.34	1.00
36	Mangrove swamp (Station 6)	0.00	0.02	0.60	1.00
37	California sublittoral	0.00	0.00	0.34	1.00
38	Lake Nyasa rocky shore	0.00	0.00	0.10	1.00
39	Lake Nyasa sandy shore	0.00	0.00	0.07	1.00
40	Malaysian rain forest	0.00	0.27	0.82	1.00

Table 3 Effect of interspecific competition on stability

Case no.	Community	Fraction of community matrices which are stable			
		$f_2=0.0$	$f_2=0.1$	$f_2=0.5$	$f_2=1.0$
1	Cochin estuary	0.12	0.10	0.09	0.03
2	Krysna estuary	0.64	0.60	0.60	0.72
3	Long Island estuary	0.49	0.39	0.41	0.33
4	California salt marsh	0.52	0.46	0.37	0.18
5	Georgia salt marsh	1.00	1.00	1.00	1.00
6	California tidal flat	0.07	0.04	0.02	0.01
7	Narragansett Bay	0.00	0.00	0.00	0.00
8	Bissel Cove marsh	0.07	0.04	0.02	0.03
9	Lough Ine rapids	0.43	0.42	0.40	0.35
10	Exposed intertidal (New England)	0.89	0.93	0.91	0.91
11	Protected intertidal (New England)	0.59	0.55	0.49	0.60
12	Exposed intertidal (Washington State)	0.41	0.36	0.35	0.42
13	Protected intertidal (Washington State)	0.19	0.28	0.48	0.35
14	Mangrove swamp (Station 1)	0.20	0.26	0.30	0.29
15	Mangrove swamp (Station 2)	0.24	0.18	0.24	0.26
16	Pamlico River	0.45	0.40	0.34	0.28
17	Marshallese reefs	0.39	0.39	0.27	0.19
18	Kapingamarangi atoll	0.33	0.28	0.01	0.00
19	Moosehead Lake	0.74	0.73	0.48	0.35
20	Antarctic pack ice zone	0.09	0.04	0.02	0.02
21	Ross Sea	0.12	0.06	0.00	0.01
22	Bear Island	0.01	0.00	0.00	0.00
23	Canadian prairie	0.19	0.17	0.09	0.04
24	Canadian willow forest	0.92	0.86	0.60	0.33
25	Canadian aspen forest	0.95	0.84	0.54	0.43
26	Aspen parkland	0.28	0.21	0.04	0.00
27	Wytham wood	0.12	0.08	0.03	0.00
28	New Zealand salt-meadow	0.78	0.58	0.15	0.03
29	Arctic seas	0.07	0.08	0.03	0.02
30	Antarctic seas	0.47	0.32	0.07	0.02
31	Black Sea epiplankton	0.00	0.00	0.00	0.00
32	Black Sea bathyplankton	0.00	0.00	0.00	0.00
33	Crocodile Creek	0.87	0.54	0.07	0.00
34	River Clydach	0.07	0.08	0.03	0.04
35	Morgan's Creek	0.71	0.72	0.38	0.29
36	Mangrove swamp (Station 6)	0.26	0.27	0.27	0.22
37	California sublittoral	0.76	0.69	0.32	0.15
38	Lake Nyasa rocky shore	0.63	0.23	0.00	0.00
39	Lake Nyasa sandy shore	0.67	0.49	0.18	0.07
40	Malaysian rain forest	0.34	0.35	0.12	0.00

argue[5] that it rarely occurs for consumer species. They seem to feel that negative A_{ii} means that intraspecific interference is not only present, but is furthermore the 'limiting factor' in the sense of Liebig. But surely any intraspecific interference, whether it is 'limiting' or not, will manifest itself in the community matrix as negative A_{ii}. For this reason I have avoided Pimm and Lawton's terminology, and feel that negative A_{ii} can be expected to occur quite often for consumer species.

The effect of interspecific interference is investigated in Table 3, with the parameter values $f_1=1$, $s_1=1$, $s_2=0.1$ fixed, while f_2, the fraction of potential competitors which participate in interspecific interference, varies. It can be seen that these interactions usually exert a destabilizing influence.

Finally, Table 4 explores the plausibility of the equilibrium viewpoint itself. The first column of this table gives the fraction of stable community matrices, constructed as above with $\{f_1, s_1, f_2, s_2\}=\{1, 1, 0, 0\}$. The second column was obtained as follows: after constructing each community matrix as above, all the positive and all the negative off-diagonal entries were randomly permuted, and then stability was checked. Thus, all predator–prey relationships were preserved, as was the average interaction strength in each matrix, but the detailed pattern of interaction strengths suggested by the particular organisms in each food web (step (1) above) was destroyed. Clearly, local stability was destroyed along with this pattern.

The third column in Table 4 results from a somewhat weaker disruption of the pattern of interaction strengths. In this case, for each community matrix constructed according to the original rules, the pairs of non-zero elements (A_{ij}, A_{ji}) were randomly permuted. Again, local stability is destroyed.

If the equilibrium viewpoint is not appropriate for these communities, there is no reason for local stability of the community matrix to be associated with a more plausible pattern of interaction strengths.

Other evidence in favour of the equilibrium viewpoint, though in a somewhat different conceptual framework from that embodied by ref. 1, has come from some simulation studies[6]. In a study which is closely related to the present one in spirit (published shortly after the submission of this paper), Lawlor[7] has found, for 11 avian and 10 reptilian competitive guilds, that

Table 4 Effect of trophic interaction pattern on stability

Case no.	Community	Fraction of community matrices which are stable 'Normal'	Disruption 1	Disruption 2
1	Cochin estuary	0.12	0.05	0.00
2	Krysna estuary	0.64	0.00	0.00
3	Long Island estuary	0.49	0.00	0.00
4	California salt marsh	0.52	0.03	0.00
5	Georgia salt marsh	1.00	0.22	1.00
6	California tidal flat	0.07	0.00	0.00
7	Narragansett Bay	0.00	0.00	0.00
8	Bissel Cove marsh	0.07	0.02	0.00
9	Lough Ine rapids	0.43	0.03	0.12
10	Exposed intertidal (New England)	0.89	0.15	0.50
11	Protected intertidal (New England)	0.59	0.12	0.16
12	Exposed intertidal (Washington State)	0.41	0.00	0.00
13	Protected intertidal (Washington State)	0.19	0.00	0.00
14	Mangrove swamp (Station 1)	0.20	0.13	0.15
15	Mangrove swamp (Station 2)	0.24	0.09	0.16
16	Pamlico River	0.45	0.00	0.12
17	Marshallese reefs	0.39	0.00	0.00
18	Kapingamarangi atoll	0.33	0.00	0.00
19	Moosehead Lake	0.74	0.00	0.00
20	Antarctic pack ice zone	0.09	0.00	0.02
21	Ross Sea	0.12	0.00	0.00
22	Bear Island	0.01	0.00	0.00
23	Canadian prairie	0.19	0.00	0.00
24	Canadian willow forest	0.92	0.04	0.02
25	Canadian aspen forest	0.95	0.02	0.04
26	Aspen parkland	0.28	0.00	0.00
27	Wytham wood	0.12	0.00	0.00
28	New Zealand salt-meadow	0.78	0.00	0.13
29	Arctic seas	0.07	0.00	0.02
30	Antarctic seas	0.47	0.00	0.00
31	Black Sea epiplankton	0.00	0.00	0.00
32	Black Sea bathyplankton	0.00	0.00	0.00
33	Crocodile Creek	0.87	0.00	0.00
34	River Clydach	0.07	0.00	0.00
35	Morgan's Creek	0.71	0.00	0.00
36	Mangrove swamp (Station 6)	0.26	0.00	0.00
37	California sublittoral	0.76	0.00	0.00
38	Lake Nyasa rocky shore	0.63	0.00	0.00
39	Lake Nyasa sandy shore	0.67	0.00	0.00
40	Malaysian rain forest	0.34	0.02	0.11

randomizations of community matrices constructed from measured resource overlaps often have a deleterious effect on stability. Although this result is suggestive, its bearing, like that of single-species studies[8], on the stability and dynamics of whole communities is not entirely clear.

Auerbach[9], applying Pimm and Lawton's community matrix rules to certain host-parasitoid systems, did not find any stable matrices. This is surely due in large part to the fact that he did not extend Pimm and Lawton's rules, as I have, to allow for varying degrees of intra- and interspecific interference. As pointed out by several authors[10-12], and reinforced by Tables 1 and 2, community matrix stability depends crucially on the extent of intraspecific interference (both in commonness of occurrence and in strength).

It has long been known that solutions to systems of autonomous ordinary differential equations can behave very strangely ('recurrent trajectories')[13,14]. More recently, the realization has grown that even very simple systems can exhibit very complicated dynamical behaviour ('chaos')[15,16]. For ecosystems, biology provides additional complications, such as seasonality, spatial migration, invasion, local extinction, fluctuating extrinsic factors, catastrophic extrinsic factors and phenotypic plasticity. In the face of all these complications, the equilibrium viewpoint is hardly to be taken for granted.

The above findings in favour of the equilibrium viewpoint legitimize a large body of ecological theory and suggest that it is worth pursuing this viewpoint further. At the same time, we should be aware of the many factors, both mathematical and biological, which complicate the picture, and strive to incorporate them into our ecological world-view.

I thank Frédéric Briand for providing me with his collection of food webs in advance of publication, and for some most interesting discussions. I also thank the University of Guelph for the provision of a substantial amount of computer time under its Special Computing Program.

Received 8 September; accepted 2 December 1980.

1. May, R. M. *Stability and Complexity in Model Ecosystems* (Princeton University Press, 1974).
2. Levins, R. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 16–50 (Harvard University Press, 1975).
3. Briand, F. *Am. Nat.* (submitted).
4. Pimm, S. L. & Lawton, J. H. *Nature* **275,** 542–544 (1978).
5. Lawton, J. H. & Pimm, S. L. *Nature* **272,** 189–190 (1978).
6. Hall, C. A. S. & Day, J. W. (eds) *Ecosystem Modelling in Theory and Practice* (Wiley, New York, 1977).
7. Lawlor, L. R. *Am. Nat.* **116,** 394–408 (1980).
8. Hassell, M. P. *J. Anim. Ecol.* **49,** 603–628 (1980).
9. Auerbach, M. J. *Nature* **279,** 821–822 (1979).
10. Saunders, P. T. *Nature* **272,** 189 (1978).
11. Pimm, S. L. *Nature* **285,** 591 (1980).
12. Rejmánek, M. & Starý, P. *Nature* **285,** 591–592 (1980).
13. Birkhoff, G. D. *Dynamical Systems* (American Mathematical Society, New York, 1927).
14. Nemytskii, V. V. & Stepanov, V. V. *Qualitative Theory of Differential Equations* (Princeton Universiy Press, 1960).
15. Lorenz, E. N. *J. atmos. Sci.* **20,** 130–141 (1963).
16. Rössler, O. E. *Z. Naturforsch.* **31a,** 259–264, 1664–1670 (1976).

Perception of sign language from an array of 27 moving spots

Vivien C. Tartter* & Kenneth C. Knowlton

Bell Laboratories, 600 Mountain Avenue, Murray Hill,
New Jersey 07974 and *Rutgers University, Camden,
New Jersey 08104

Many deaf people in the USA communicate in American sign language (ASL), which has an expressive capacity equivalent to that of a spoken language, although structurally independent of spoken languages[1,2]. It comprises hand and arm movements often combined with particular facial gestures; together these are sufficiently precise to transmit all the complexities and innuendoes of a language. Here we demonstrate that fluent ASL users can communicate easily when all they see of each other is an array of 27 light spots strategically placed on the hands and face. The results indicate the salient locations in normal sign perception, and suggest that it is feasible to transmit signs using the bandwidth of one telephone line rather than a much more expensive TV line.

In normal conditions, the understanding of sign language requires seeing the area in front of the signer's body from the top of the head to waist level, and within a few inches of each side—signs produced within a 6-inch radius of the chin are most precisely articulated[3]. Signs are differentiated from one another by hand shape (for example, index finger compared with little finger raised), hand orientation (for example, palm exposed rather than back of the hand exposed), type of movement (for example, movement towards compared with movement away from the signer), and 'place of articulation' (for example, a touch on the chin compared with a touch on the lips)[4,5]. Battison[5] estimated that there are 45 distinct hand shapes, 10 hand orientations, 10 movements and 25 locations—this indicates the precision required to distinguish signs. In addition it has been

0028-0836/81/080676—03$01.00

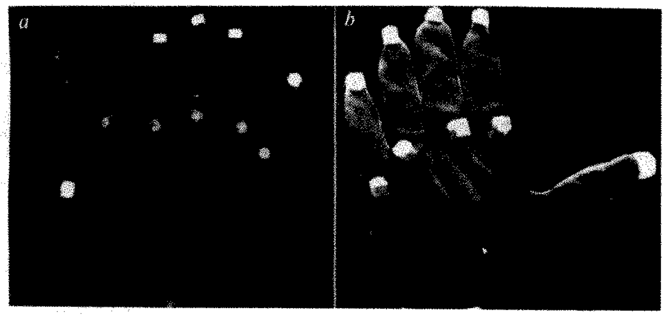


Fig. 1 A display of the back of a glove in *a*, lighting conditions in which only the reflections from the tape are visible and *b*, in normal lighting conditions.

suggested that facial expression may transmit critical information, perhaps redundantly[5], and that head movements may be critical for differentiating affirmatives, negatives and interrogatives[6].

To determine which information is critical for evoking the perception of the various features, we adapted the method of Johansson *et al.*[7,8] to study grosser types of biological motion. Johansson attached patches of retroreflective tape to the principal joints of subjects—wrist, elbow, shoulder, ankle, knee and hip—then filmed them moving in special lighting conditions so that only the patches of tape were visible. The two-dimensional display of light spots produced was clearly and immediately identified by observers as human motion: a person walking, running, cycling, and such. Presentation of a still picture alone from the display resulted in perception of lights—the motion was necessary to convey the structural relation between the lights and thus to evoke perception of human action.

As signing comprises human movement, although faster, more finely detailed, and less repetitive than that studied by Johansson, it seemed possible that the same perceptual mechanisms which organized the spots of light in Johansson's displays into meaningful patterns could organize strategically placed light spots on the hands into hand shapes and motions. The problem was to derive a configuration of spots which would convey each of the important feature values in sign without confusion, and which would be immediately perceived as moving hands, suggesting that these points are normally the salient ones for sign perception. (Note that Poizner *et al.*[9] used a similar technique to code the motion patterns used for various inflections in ASL. For this purpose one light on the index fingertip and reference points are sufficient to distinguish between 50 movement types.)

We devised a glove with 13 strategically placed pieces of tape, the largest and brightest at the regions which convey most information: a large piece on the thumb tip and slightly smaller ones on each of the remaining fingertips. On the back of the hand there was a smaller spot on each finger, just proximal to the second joint from the tip and there were four even smaller spots around the wrist. The backs of two gloves are shown in Fig. 1 in lighting which shows only the retroreflective tape spots (*a*), and in normal lighting (*b*).

With the palm of the hand forward, an outline of the hand is perceived (from the fingertips and wrist boundary). There is no hand shape which can be assumed with the palm forward in which the fingertips are not visible—closing of the hand is evident from the motion and the approach of the tip spots towards the wrist. Elevation of individual fingers is clear from the position of the raised fingers relative to the folded ones.

The back of the hand appears as fingertips followed by two more rows of spots. The row at the joints is necessary because with the back of the hand forward it is possible to curve the fingers so that the tips are not visible, as in a fist. In this case, curvature of the fingers is apparent as a disappearance of the fingertip spots, but information on hand shape is still preserved by the relative positions of the joint spots.

Information about hand orientation is given again by motion and the configuration of spots. Each spot is actually a spherical protrusion which is therefore uniformly visible from a wide range of viewing angles. Motion information is readily apparent in horizontal and vertical directions, as it was for subjects used in other studies[8,9], from the changes in position of the spots. Motion in depth is conveyed by a change in the size and spacing of the spots due to foreshortening.

To convey place of articulation only one additional reference point was necessary—a spot on the nose. A touch at the eyebrows is seen as a hand motion above the middle of the face; at the shoulders, as outside and below the middle of the face. No further reference points on the body were necessary, perhaps because of the grosser articulation outside the face area. The spot on the nose, in addition to providing a reference point, also gave information on head movements (nodding, shaking and

S_1	S_2
	How is your work?
[confused, so shadows what saw] How is . . .	
	How is your work?
[now understands] OK, OK.	
	Same, same. Boring
Same, same. OK	
	Thinking new job, I mean transfer to another job? Still waiting for another job? You. You!
For you? me. Thinking about transferring to another job maybe.	
	Same place or other place?
Same. Here. But change job.	
	For how long have you been waiting?
No. No. No. Been thinking for a pretty long time to get transferred so I have to wait for my supervisor to help me get another job.	
	Are you going to night school this fall?
I'm thinking of writing to Middlesex County College for night college but I have to wait for the reply. Still waiting for a man who is responsible for picking the deaf people to put into one class. People.	
	It seems to me that you have been thinking for a very long time. You'd better stop thinking and do it now, before it becomes too late. OK? [Pauses with no response] Do you understand me?
Yes. Talk. That I'm not worried about it.	
	Noooo . . ! You have been concerned for a *long* time. Concerned. Not important to think. Do it before it becomes too late.
Oh. I understand. You mean that if I . . .	
	You wait too long. No don't. Waiting is not that good.
I understand.	
Today it's very hard to get into college. Very difficult. Because too many people want enrolment.	No. No.
	No. Do you know what law number 504 means to you? Always assume that it's not hard to get in because you are deaf. No. Enrol (any) time you want. 504 guarantees enrolment.
[S gives feedback] Yes. Yes. I get it. I understand.	
But how can I get in? I mean I have to get permission from someone to get in.	
	Get. You want to get into college you have to go to admissions office and ask what courses you want; then he will put down (what) you ask him. You pay an interpreter, you pay for, and if college says no, you bring 504 law and show them, and if college won't believe you go to a court and ask for some -- [interrupts herself and restarts immediately with] because I learned from my friend who refused - (I mean) college refused to pay nothing so she went to court and the lawyer went to the college and college realized their mistake and they will pay for all interpreters. (It's) easy. You don't have to pay nothing.
Oh. I understand.	[Pauses briefly, then teases with] Chicken, chicken.

Fig. 2 One conversation of 4 min transcribed into English from videotaped record. Subjects at this point had 10 min experience with the system. Average signing rate was ~80 signs per min. S_1, S_2, subjects 1 and 2 respectively.

leaning forward), and provided feedback information to the person signing that the viewer was still watching.

Preliminary tests showed that effective communication at near normal rates is possible when viewing only the 27 points of light. Two pairs of deaf subjects were able to converse when viewing only the other's spots. Three of the four subjects were profoundly deaf, with hearing losses of 90 dB in the better ear. The fourth subject (C.F.) had some hearing and quite good speech. She and her partner were well educated with good reading and writing skills. Both learned ASL when young, and sign either ASL or English. The other two subjects had poor reading and writing skills and were only able to sign ASL. The transcript shown in Fig. 2 is an extract from the conversation of the first pair, translated into 'English' by C.F.

Each member of the pair was seated in a private room ~3 feet from a 21-inch TV monitor and a camera. The gloves and nose spot were placed on the subject, who was allowed a few minutes of self-viewing in normal lighting conditions. Then all the lights in each room except the one next to the TV camera were turned out, and the brightness and contrast of the TV monitors adjusted so that each subject could see only the spots. (Note that the camera was set so that movement within the signing space would be visible; thus the screen image of the spots was smaller than that displayed in Fig. 1. In addition, with movement, the spots commonly smear, leaving a slightly blurred image on each frame.) After self-viewing, the equipment was reconnected so that the subjects could see each other. They were instructed to discuss anything they wished and a free-flowing 20-min conversation ensued. Various topics were discussed by both pairs, with difficulties and requests for repetition only in finger spelling. A transcription of a portion of one pair's discussion is shown in Fig. 2. During this displayed segment the participants became quite heated—emotion was conveyed by the slow and abrupt movements of the spots. Nonmeaningful movements (for example, tugging at the gloves) were occasionally mimicked playfully, but no attempt was made to interpret them. Note that during debriefing both pairs of subjects reported that they really needed to concentrate, but that their conversation and understanding were normal except for finger spelling. Furthermore, all four subjects expressed their enjoyment of the system.

We believe that the picture that these subjects used could be coded to the capacity of one telephone line. This assumes that the useful information in each spot is its brightness or size, and position on the screen in time. A more controlled test is planned, using computer processing to ensure that this information alone is available.

We have thus shown that communication in sign language is possible using a highly stylized representation of the hands and body, at a considerably reduced information rate from a face-to-face situation. This demonstration suggests the possibility of developing a sign language telephone.

We acknowledge the assistance of Jay Jackson, Carol Furino, Cynthia Jefferey, Michael Androsko and Gene Marshall in testing our system. We also thank Lilah Axelsen for teaching us sign language, Eric Albert and Carol Furino for help in interpreting tapes, Carol Tipton and Susan Fischer for consultation about deaf communication and sign language, and Bela Julesz and Osamu Fujimura for helpful suggestions. A portion of this paper was presented at the Nato Advanced Study Institute: Sign Language Research, in Copenhagen, August 1979.

Received 22 July; accepted 18 December 1980.

1. Bellugi, U. & Klima, E. S. in *Origin and Evolution of Language and Speech* (eds Harnad, S. R., Steklis, H. D. & Lancaster, J.) 515–538 (N.Y. Acad. Sci., 1976).
2. Grosjean, F. in *Recent Perspectives in American Sign Language* (eds Lane, H. & Grosjean, F.) (Lawrence Erlbaum, New Jersey, in the press).
3. Siple, P. in *Understanding Sign Language through Sign Language Research* 3–23 (Academic, New York, 1978).
4. Stokoe, W. C., Casterline, D. & Croneburg, C. *A Dictionary of American Sign Language on Linguistic Principles* (Gallaudet College Press, Washington DC, 1965).
5. Battison, R. *Lexical Borrowing in American Sign Language* (Linstok, Maryland, 1978).
6. Bellugi, U. & Fischer, S. *Cognition* **1,** 173–200 (1972).
7. Johansson, G. *Percept. Psychophys.* **14,** 201–211 (1973).
8. Johansson, G., von Hofsten, C. Jr & Jansson, G. *A. Rev. Psychol.* 27–63 (1980).
9. Poizner, H., Bellugi, U. & Lutes-Driscoll, V. *Journal of Experimental Psychology: Human Perception and Performance* (in the press).

Simultaneous yet independent inheritance of somatically acquired tolerance to two distinct H–2 antigenic haplotype determinants in mice

R. M. Gorczynski

Ontario Cancer Institute, 500 Sherbourne Street, Toronto, Ontario, Canada M4X 1K9

E. J. Steele

Transplantation Biology Section, Clinical Research Centre, Harrow, Middlesex HA1 3UJ, UK

We have previously presented evidence that somatically acquired immunological tolerance to foreign histocompatibility antigens induced in male inbred mice can be transmitted at a high frequency to first- and second-generation offspring without exposure of the progeny to the tolerizing antigenic stimulus[1]. These data have in turn been interpreted to support a hypothesis predicting soma to germ-line inheritance for acquired states of the immune system[2]. Formal proof of this genetic scheme at the molecular level is not yet available. However, an even stronger reason for abandoning the notion of the isolation of the germ line from the soma (Weismann's doctrine[2]) would be provided by a clear demonstration of the multiple inheritance of independently acquired somatic characters. Here we present evidence that individual male mice (B10) made tolerant to two different H–2 haplotypes (B10.BR, B10.D$_2$) can transmit the tolerance state to each haplotype independently, yet often simultaneously, at a high frequency to both first and second generation progeny.

Congeneic resistant mouse lines (B10.SgN, $H\text{–}2K^bD^b$; B10.D$_2$, $H\text{–}2K^dD^d$; B10.BR, $H\text{–}2K^kD^k$) were purchased at 6 weeks of age and ATH mice ($H\text{–}2K^sD^d$) at 8 weeks of age from Jackson Laboratories. The latter mice were used within 4 weeks of arrival. Normal litters of B10.SgN (hereafter called B10) were produced in our animal room from the Jackson foundation stock; some members of these litters were mated (brother × sister) to produce normal B10 animals on a regular basis during the 10-month study. F$_1$ hybrids (B10 × B10.D$_2$, B10 × B10.BR) were produced by matings from the Jackson foundation stock. These served as a source of lymphoid cells both for inducing and maintaining tolerance in the B10 mice, and as normal responder or stimulator cells in the cytotoxic T-cell assays.

Tolerance was induced in neonatal B10 mice (produced at the start of the breeding programme) by intraperitoneal (i.p.) inoculation of 100×10^6 F$_1$ lymphoid cells (approximately 1:1 mixtures of bone marrow and spleen cells, Fig. 1; bone marrow only, Fig. 2). As previously reported, repeated injections of the same type of F$_1$ cells were given at 2-week intervals throughout the study (i.p. injections until 4 weeks of age—thereafter intravenous injections were given). While certainly not essential to produce operational tolerance in such neonatally challenged mice[3,4], we have continued these regular inoculations of F$_1$ hybrid cells, as discussed earlier, to ensure high levels of circulating F$_1$ hybrid cells in the neonates. No data are yet available to judge whether such repeated challenge is essential to induce the transmission of tolerance from parents to offspring which we have shown subsequently occurs (ref. 1 and below).

Tolerant male mice were chosen for mating with normal B10 females beginning at 8 weeks of age. Tolerance in these breeding fathers has been assessed by both cell-mediated lympholysis (CML) responsiveness (performed after partial splenectomy under ether anaesthesia) and skin graft rejection (see legends to Figs 1, 2). Only animals judged to be tolerant by both criteria

0028–0836/81/080678—04$01.00

have been used in the breeding studies reported. However, responsiveness in the progeny of such animals has been assessed initially only by *in vitro* assay (see below). Comparison of skin grafting and CML responses of the progeny is being investigated. First-generation progeny of these animals were reared without exposure to F$_1$ antigens. At 8 weeks, spleen cells of individual progeny mice were assayed *in vitro* for their ability to give rise to cytotoxic lymphocytes after stimulation with irradiated (B10 × B10.BR)F$_1$, (B10 × B10.D$_2$)F$_1$, or third-party ATH (H–$2K^sD^d$) spleen cells (CML assay). Target cells in the ^{51}Cr-release test were concanavalin A blasts of the respective spleen cells used as stimulators[4–6].

Typical data for first-generation progeny of B10 male mice made tolerant to B10.BR or B10.D$_2$ determinants are shown in Fig. 1. Clearly these first-generation animals show a high frequency (some 40–50%) of tolerance (0–5% specific cytotoxicity) to those determinants used to induce tolerance in their fathers (B10.BR, Fig. 1 *d*; B10.D$_2$, Fig. 1*e*). Note that not only is the ability of the progeny to respond to the third-party antigens (ATH) unaffected but so too is their ability to respond to B10.D$_2$ antigens (progeny of males tolerant to B10.BR) or B10.BR antigens (progeny of males tolerant of B10.D$_2$). It is also clear (Fig. 1*f*) that first-generation mice born to males undergoing the tolerance-inducing treatment procedure, but which failed to become tolerant, are phenotypically (in terms of CML responses) equivalent to the offspring of normal B10 mice. These results confirm our earlier findings on the genetic transmission of acquired tolerance to A-strain H–2 antigens in CBA mice[1], and extend the generality of the phenomenon to two additional mouse strain combinations.

Clearly acquired tolerance to B10.D$_2$ antigens can be inherited independently of tolerance to B10.BR antigens (Fig. 1). Interestingly, an apparent simultaneous transmission of tolerance to the independent K/D-end specificities associated with a given haplotype (for example, H–2^k) has occurred, as these specificities are themselves independently expressed on separate molecules. Although this in itself may represent the simultaneous transmission of tolerance to independent antigenic determinants, note that responsiveness to the distinct K/D specificities was not tested independently in the recipients described here, and alloreactive cytotoxicity directed to K end-encoded determinants is more readily demonstrated than to D end-encoded determinants. Thus one interpretation of our data would be that responsiveness to major stimulating antigen determinants (K end-encoded) only was recorded in the CML assays used and there is no need to postulate the parallel transmission of immunological tolerance to K^k, D^k in animals bred from males tolerant to B10.BR cells. Indeed, recent studies (unpublished) indicate the independent transmission of tolerance to K^k, D^d from B10 male mice made tolerant to B10.A lymphoid tissue and subsequently mated with normal B10 female mice.

We have investigated whether tolerance to two different specificities can be independently transmitted to progeny using a situation in which individual B10 males were rendered tolerant simultaneously to both H–2^k and H–2^d haplotypes. Hereafter we will make the minimum interpretation that we are measuring unresponsiveness to the major stimulating specificities (K^k, K^d) associated with those haplotypes. The data in Fig. 2 summarize results from first-generation progeny (of four such males) obtained over a 12-week breeding period. The data are arranged as response profiles to B10.D$_2$ or ATH antigens for progeny mice grouped into different levels of responsiveness (% specific cytotoxicity) to B10.BR antigens. (A similar plot is obtained if the progeny are grouped into different levels of response to B10.D$_2$.) Note: (1) All progeny mice respond to ATH antigens, irrespective of their response patterns to each of the other two antigens tested. (2) The response pattern to one haplotype is not predictive of the level of responsiveness obtained towards the other haplotype (for example, of 80 progeny hyporesponsive to B10.BR—$<10\%$ cytotoxicity—40 respond in the normal range to B10.D$_2$ and 40 are tolerant or low responders to B10.D$_2$). That is, in contrast to their fathers which are tolerant of both haplotypes, many of these first-generation offspring are now tolerant for either one or the other haplotype. (3) Many of these progeny also score tolerant (27/150) to or are low responders (13/150) of both H–2 antigenic types. These results therefore show that the inheritance of

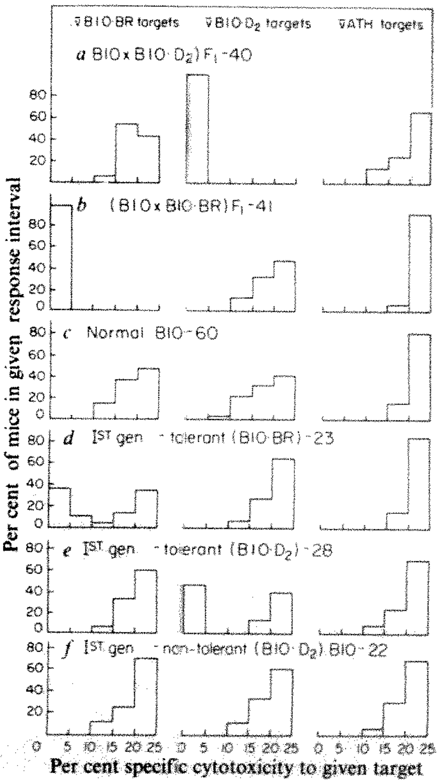


Fig. 1 CML response profiles of first-generation offspring of male B10 mice rendered tolerant to B10.BR or B10.D$_2$ antigenic determinants. Spleen cells (1.5×10^6 in duplicate) from individual 8-week progeny mice or control mice were stimulated in culture with irradiated (1,500 rad) spleen cells (5×10^5) of the relevant targets shown. Cytotoxicity was assayed at 5 days (refs 4–7). *a*, *b*, *c*, Responses of normal control mice (B10 × B10.D$_2$)F$_1$, (B10 × B10.BR)F$_1$ and B10 respectively. (The number of mice tested is indicated.) *d*,*e*, First-generation progeny of male B10, rendered neonatally tolerant to (B10 × B10.BR)F$_1$ (*d*) or (B10 × B10.D$_2$)F$_1$ lymphoid cells (*e*), mated to normal B10 females. All tolerant males used for breeding were specifically unresponsive to the tolerizing antigenic determinants in 5-day CML tests (performed in spleen lymphocytes obtained by partial splenectomy under ether anaesthesia) and retained a skin graft of the appropriate F$_1$ hybrid for at least 100 days (while rejecting third-party grafts in the expected fashion). No data from first-generation progeny born to animals not tolerant by these criteria are shown. Note that the efficiency of inducing stable tolerance in the present strain combinations was lower than in the previous study using (CBA × A)F$_1$ → CBA (ref. 1). Thus, of 9 male and 4 female B10 mice injected at birth with (B10 × B10.D$_2$)F$_1$, 4 males and 2 females scored tolerant by CML tests and 4 males and 0 females scored tolerant by skin grafting. Of 8 male and 5 female B10 injected at birth with (B10 × B10.BR)F$_1$, 4 males and 1 female were tolerant by CML and 4 males and 0 females were tolerant by skin grafting. Data shown represent cumulative data for progeny of all the tolerant (normal) fathers of a given type (at least 1 litter per father; minimum litter size, 4 offspring). Normal B10 mice tested were age-matched with tolerant progeny and were produced from mating seven normal male B10 mice with normal B10 femals. *f* Shows a limited study (22 mice) of the offspring to B10 mice inoculated neonatally with (B10 × B10.D$_2$)F$_1$ cells but shown thereafter (by CML and skin graft testing) to be fully responsive to B10.D$_2$ alloantigens. Offspring from all five males are included in this study (minimum litter size, 3). Cytotoxic data for individual mice were derived from 4-h ^{51}Cr release tests with an effector : target cell ratio of 2 : 1. total releasable c.p.m. and spontaneously released c.p.m. (arithmetic mean ± s.e.m.) for the different targets used were: B10.BR: 9,161 ± 139, 633 ± 11; B10.D$_2$: 10,132 ± 129, 931 ± 17; ATH: 8,149 ± 101, 576 ± 10. Assays for all targets and all effector populations were performed in triplicate. To test for significance between responses of normal B10 mice and progeny, response groups were compared using 2×2 or $2 \times 3 \chi^2$ tables. (For 2×2, responses grouped as $\leq 10\%$, $>10\%$ or $\leq 20\%$, $>20\%$ specific cytotoxicity; for 2×3, $\leq 15\%$, 15–20%, $>20\%$ specific cytotoxicity.) χ^2 and probability values are: response to B10.BR: *c* versus *d*, 29.7, $P < 0.001$; *c* versus *e*, 2.8, $P > 0.2$; response to B10.D$_2$: *c* versus *d*, 4.4, $P > 0.1$; *c* versus *e*, 28.6, $P < 0.001$; response to ATH: *c* versus *d*, 0.048, $P > 0.8$; *c* versus *e*, 1.2, $P > 0.2$.

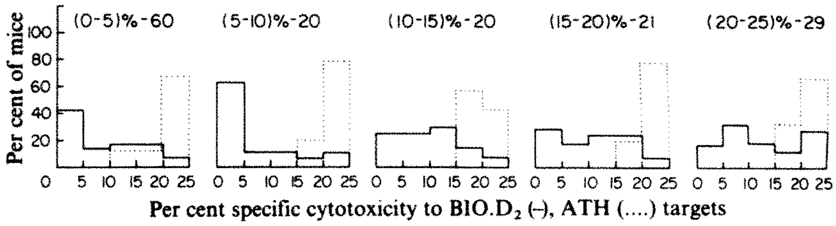


Fig. 2 CML response profile of first-generation progeny born to B10 males tolerized as neonates to both (B10×B10.D$_2$)F$_1$ and (B10×B10.BR)F$_1$ lymphoid cells. In this case B10 neonates were inoculated intraperitoneally with a 1 : 1 mixture of 50×10^6 bone marrow cells from the respective F$_1$ hybrid donors, and these inoculations were repeated every 2 weeks throughout the study. Spleen cells from 8-week-old progeny of matings of B10 animals known to be tolerant to both haplotypes (acceptance of skin grafts from both parents; unresponsiveness in CML cultures of spleen cells obtained by partial splenectomy) were stimulated in culture with irradiated (B10×B10.BR)F$_1$, (B10×B10.D$_2$)F$_1$ or ATH spleen cells and cytotoxic responses determined 5 days later. Individuals were then grouped according to their level of cytotoxic response to one tolerizing antigen (B10.BR) and the per cent of mice responding with a given degree of cytotoxicity (0–5%, 5–10% and so on) to the other two antigens recorded. Note (see Fig. 1 legend) that of 19 males and 8 females injected at birth with the bone marrow cell mixture, 6 males and 4 females were tolerant to both B10.D$_2$ and B10.BR in CML assays, and 4 males and 2 females were tolerant to both by skin grafting. As described in Fig. 1 legend, offspring from all 4 males are included in the data shown here (at least 1 litter per male; minimum litter size, 4 offspring). As in Fig. 1, cytotoxic assays were performed in triplicate at an effector: target cell ratio of 2 : 1. Total releasable c.p.m. and spontaneously released c.p.m. (arithmetic mean ± s.e.m.) for the targets shown were: B10.BR: 8,104 ± 106, 609 ± 11; B10.D$_2$: 9,112 ± 113, 613 ± 13; ATH: 7,048 ± 96, 596 ± 14. To test the null hypothesis that the response to B10.D$_2$ is independent of the response to B10.BR, a 3 × 3 table was constructed (response levels being ≤5%, 5–15%, >15% specific cytotoxicity). The $\chi^2 = 8.4$ and $0.05 < P < 0.1$ (null hypothesis cannot be rejected). The group significantly lower than expected (disturbing the rest of the table) was that group responding high to B10.BR (>15%) and low to B10.D$_2$ (≤5%), which contained fewer animals (11) than expected (18.7).

somatically acquired tolerance to two different H-2-encoded haplotypes is assorted at random in the first generation but that a high percentage show evidence for simultaneous transmission of these two specific tolerant traits.

Finally we examined the effect of breeding from the first-generation offspring—derived originally from male B10 mice tolerant to both B10.D$_2$ and B10.BR—on the subsequent transmission of the tolerant traits to second-generation offspring (in all cases only the original grandfathers were ever deliberately exposed to the tolerizing antigenic determinants). Three types of first-generation progeny were used as breeders: those which scored individually tolerant to B10.BR or B10.D$_2$ determinants (Fig. 3*a*, *b*) and those tolerant to both determinants (Fig. 3*c*). These assignments of parental phenotype were made on the basis of CML assays with spleen cells obtained by partial splenectomy.

The data of Fig. 3*a*, *b*, although limited to studies of only 27 and 15 individuals respectively, are clearly quite different from those of Fig. 2, and are, in fact, similar to those obtained by re-plotting the data of Fig. 1*d*, *e* in this format. Thus, first-generation progeny parents phenotypically tolerant to B10.BR give rise to second-generation offspring which respond to B10.D$_2$ antigens, yet which have retained their unresponsive phenotype for B10.BR antigens (Fig. 3*a*). Similar effects are also seen in the second-generation progeny derived from first-generation parents tolerant to B10.D$_2$ antigens (Fig. 3*b*). No unexpected reacquisition of tolerance for the other MHC haplotype was observed in any of the 42 mice tested (see below). In contrast to these patterns, first-generation progeny breeders scoring tolerant to both B10.BR and B10.D$_2$ antigens give rise to second-generation progeny (Fig. 3*c*) with the same 'spray' of cytotoxic response patterns as seen with the initial first-generation progeny (compare with Fig. 2). Some animals were tolerant either to B10.BR or B10.D$_2$ antigens, some were still tolerant to both antigens, and some had lost tolerance to both antigens.

In summary, investigation of the inheritance of neonatally acquired tolerance to major histocompatibility complex (MHC) gene products in congeneic resistant mice (using a B10 background) has shown that two separately determined tolerance phenotypes can be transmitted and assorted independently in first-generation progeny derived from matings of males tolerized to two H-2 haplotypes with normal females; and moreover, that a significant proportion of these first-generation progeny are also unresponsive towards determinants encoded by genes associated with both haplotypes. Transmission of these unresponsive traits is relatively stable as indicated by the fact that no spontaneous reversion (to the 'double' tolerance state) was observed in second-generation progeny which failed to display tolerance to one of the two H-2 haplotypes, and a significant percentage (16%, $P < 0.05$; see left-hand panels of Fig. 3*c*) of second-generation mice born to parents tolerant to both haplotypes still fail to respond to both, although independent assortment of the two responses can also be observed. These results suggest that genetic transmission of acquired tolerance to H-2 antigens may assort in a mendelian manner.

We noted what may be a more apparent than real dichotomy in the independent transmission of tolerance to $H\text{-}2^k$, $H\text{-}2^d$ haplotypes (from mice tolerant of B10.A), yet the simultaneous transmission of unresponsiveness to $H\text{-}2K^k$, $H\text{-}2D^k$ from mice tolerant of B10.BR. As we pointed out, we have not rigorously tested the response of mice to either K^k or D^k determinants (although independent transmission of tolerance to K^k, D^d does occur from B10 mice tolerant to B10.A MHC determinants, unpublished results). We reported previously a phenomenon of 'skipping' of expression of the tolerance phenotype in sequential generations of CBA mice made tolerant to (CBA×A)F$_1$ antigens[1]. Thus second-generation tolerant progeny are derived from apparently normal responder first-generation mice. In contrast, we recorded (Fig. 3) here a failure to detect 'reversion' or skipping) in sequential generations of mice. We cannot yet explain these differences, but it is not unreasonable to anticipate that they are due to inherent differences in the two experimental situations: for instance, to the ease with which tolerance can be produced in CBA (to A/J antigen) relative to the difficulty reported here in tolerizing B10 mice.

Finally, note the apparent difference in the behaviour (as assessed by CML response) of progeny mice derived from mating A-strain females with either A males neonatally tolerized to (A×B)F$_1$ cells or normal (A×B)F$_1$ males. In the latter case (A × A), offspring are not unresponsive to B antigens, despite the fact that tolerance to both A and B in F$_1$ mice presumably takes place after fusion of gametes. Nevertheless, the data presented here and elsewhere show that in the former situation the progeny acquire the tolerance phenotype of their parents. We cannot yet explain convincingly this anomaly; however, there are apparently many mechanisms whereby somatically acquired tolerance may be induced and maintained[7], and so far there is little evidence to indicate which, if any, of these mechanisms represents the natural mode of self/non-self discrimination.

The high frequency of parallel transmission of independently assorted, somatically acquired characters severely challenges Weismann's doctrine, and seems to us to demand an explanation not offered by conventional neo-darwinian genetics. It has been argued[2] that the type of phenomenon documented here and elsewhere[1,2,8,9], may be explained best by invoking the clonal selection and somatic mutation theory of immunology[10-13] and soma germ-line transmission of somatic genetic information (RNA) through endogenous RNA viral vectors and reverse transcriptase[14,15]. We might then imagine that genetic transfer of somatically selected information associated with the regulation of responses, for example, to A/J antigens in CBA mice[1] (anti-idiotype ?, antigen ?) is crucial for the anomalous inheritance patterns observed.

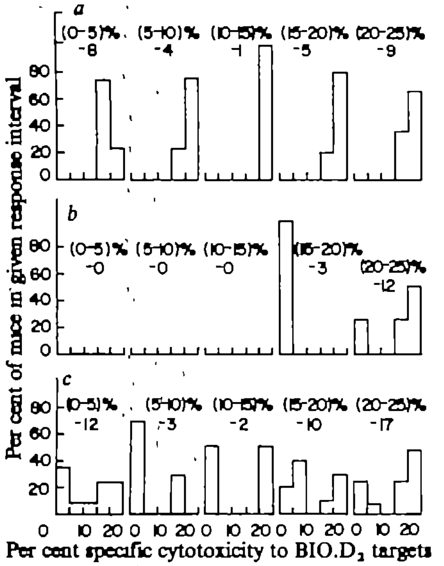


Fig. 3 Analysis of transmission of response patterns from first-generation progeny (Fig. 2) to second-generation progeny. Three types of first-generation progeny were used as parents to produce this data. In *a*, animals were selected (by CML tests performed after partial splenectomy) which were tolerant only to B10.BR stimulator cells—these first-generation progeny gave normal responses after stimulation with (B10×B10.D_2)F_1 cells. *b* Shows second-generation progeny derived from first-generation animals found to be tolerant only to B10.D_2 antigens. The data of *c* were derived from second-generation progeny born to animals found to have inherited the tolerance phenotype of the original male parents unresponsive to both B10.D_2 and B10.BR antigens). In all cases breeding was performed by mating either phenotypically tolerant males/females to normal females/males (outcross) or by incrossing of first-generation progeny with identical CML response phenotypes (earlier studies have suggested little difference in the gross inheritance of tolerance from such matings)[1]. Thus, for *a* the data are pooled from 5 litters from incross matings and 1 litter (5 progeny) of an outcross first-generation female; for *b* the data are pooled from 3 litters of 3 incross matings, while for *c* the data represent cumulative studies on 10 litters (≥1 per cage; minimum size, 4 mice) of 4 incross matings and 3 outcrosses (using first-generation tolerant female). The data of these second-generation progeny (total individual numbers tested for *a*, *b*, *c*, were 27, 15, 44 respectively) were grouped into response intervals for one antigen and the frequency of response in a given cytotoxic interval for the other antigen plotted (see Fig. 2). As in Figs 1, 2, cytotoxic responses after stimulation with ATH cells were also determined, but the pattern of response was not qualitatively or quantitatively different from that shown in Figs 1, 2 (all animals responded, irrespective of response status to B10.D_2 or B10.BR, data not shown). All data were obtained from the assays performed in triplicate at an effector:target cell ratio of 2:1. Total releasable c.p.m. and spontaneously released c.p.m. (arithmetic mean ± s.e.m.) for the different targets used were: B10.BR: 9,100±113, 689±16; B10.D_2: 7,068±94, 495±12; ATH: 10,351±129, 689±21. To test the null hypothesis that the response to B10.D_2 is independent of the response to B10.BR (*c*), a 2×2 table was constructed (≤15%, >15%) and an exact χ^2 test performed giving $P > 0.2$ (null hypothesis cannot be rejected).

We thank Susannah MacRae, Barbara Khomasurya, Susan Oliphant and Sue Bejasa for assistance. For critical comments on the manuscript we thank Drs Elizabeth Simpson, Bob Blanden, Arno Mullbacher and Jeff Pollard; and Walter Fierz and Steve Molnar for advice on statistics. This work was supported by the MRC of Canada (grant MA 5440) and the NCI of Canada. E.J.S. thanks the KIB Foundation (UK) for support.

Received 11 June, accepted 3 December 1980.

1 Gorczynski, R. M. & Steele, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2871–2875 (1980).
2 Steele, E. J. *Somatic Selection and Adaptive Evolution* (Williams & Wallace, Toronto, Croom Helm, London, 1979).
3 Billingham, R. E., Brent, L. & Medawar, P. B. *Phil. Trans. R. Soc.* **B239**, 357 (1956).
4. Gorczynski, R. M., MacRae, S. & Till, J. E. *Scand. J. Immun.* **7**, 453–465 (1978).
5. Gorczynski, R. M. & MacRae, S. *J. Immun.* **122**, 737–746 (1979).
6 Gorczynski, R. M. *Immunology* **33**, 697–712 (1977).
7 Gorczynski, R. M., Khomasurya, B., Short, L. & MacRae, S. in *Regulation by T Cells* 18 (University of British Columbia Press, 1979).
8 Guttman, R. D. & Aust, J. B. *Nature* **197**, 1220–1221 (1963).
9 Kanazawa, R. D. & Imai, A. *Jap. J. exp. Med.* **44**, 227–234 (1974).
10. Burnet, F. M. *The Clonal Selection Theory of Acquired Immunity* (Cambridge University Press, London, 1959).
11 Cohn, M. *Progress in Immunology II* Vol. 1 (eds Brent, L. & Holborow, J.) 261–284 (North Holland, Amsterdam, 1974).
12 Jerne, N. K. *Eur. J. Immun.* **1**, 1–9 (1971).
13 Cunningham, A. J. *Contemp. Topics molec. Immun.* **3**, 1–26 (1974).
14 Temin, H. M. *J. natn. Cancer Inst.* **46**, 3–8 (1971).
15. Temin, H. M. *Science* **192**, 1075–1080 (1976).

B220: a B cell-specific member of the T200 glycoprotein family

Robert L. Coffman & Irving L. Weissman

Laboratory of Experimental Oncology, Department of Pathology,
Stanford University Medical Center, Stanford, California 94305

T200, a major cell-surface glycoprotein on lymphoid cells, exists in several forms with different electrophoretic mobilities. These forms have been correlated with different classes of lymphoid cell. The smaller forms, with molecular weights (MWs) of ≃170,000 and 180,000, are found predominantly on T cells while the 220,000 MW form is associated with B cells[1,2]. The polypeptide portions of each molecule may be identical or closely related as all three forms share the same allelic variations[3-5], and all reported heterologous antisera and monoclonal antibodies to T200 precipitate all three forms[2-6]. We report here a monoclonal antibody specific for the 220,000 MW form of T200 and show that it is expressed only on B cells and a subset of bone marrow cells which includes B cell precursors. We suggest that this form of the molecule be designated provisionally B220.

The hybridoma RA3-3A1 was produced by the fusion of the mouse myeloma S194/5.XX0.BU-1 with spleen cells from a rat immunized with the mouse cell line RAW 112, a pre-B cell tumour derived by infection of a mouse with the Abelson murine leukaemia virus[7]. The monoclonal antibody-secreting hybridoma RA3-3A1 was selected from the other fusion products because it produced an antibody that recognized RAW 112 cells and a subset of bone marrow cells, but not thymocytes. About 30% of bone marrow cells, 50% of spleen cells, 40% of mesenteric lymph node cells and less than 2% of thymocytes were positive when stained with the optimum titre of RA3-3A1 followed by fluorescein-conjugated rabbit anti-rat immunoglobulin. Flow cytometry analysis was carried out using the fluorescence-activated cell sorter (FACS III, Becton Dickinson, California). RA3-3A1 is an IgM and binds directly to the bacterium *Staphylococcus aureus*.

Immunoprecipitation of surface-iodinated spleen and RAW 112 cell lysates followed by SDS-polyacrylamide gel electrophoresis (PAGE) showed that RA3-3A1 precipitated a molecule with an apparent MW of 220,000 (Fig. 1). Precipitation of the same lysates with a monoclonal anti-T200 antibody gave the expected three bands from the spleen, the

0028-0836/81/080681—03$01.00

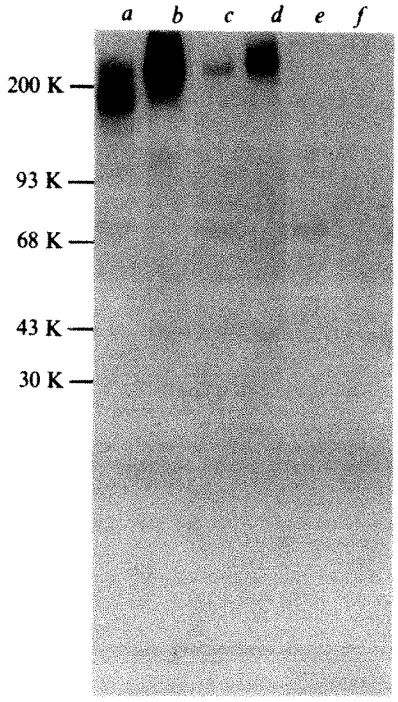


Fig. 1 Immunoprecipitation with RA3-3A1 and anti-T200. BALB/c spleen cells and RAW 112 cells were surface-iodinated by the lactoperoxidase method[8]. The cells were lysed with 1% NP40 in 0.01 M sodium phosphate buffer, *p*H 7.5, and the lysates clarified at 100,000*g* for 1 h. Aliquots of lysate from 2×10^6 cells were incubated with monoclonal antibodies and then precipitated with *S. aureus*[9] (for RA3-3A1 and control IgM) or *S. aureus* saturated with purified rabbit anti-rat Ig antibodies (for anti-T200). The rat anti-T200, prepared from the hybridoma 55-10.2 (ref. 10), was given by Dr J. A. Ledbetter. Control precipitates were prepared from rat IgM monoclonal antibody with no apparent reactivity to mouse cells. Immune precipitates were washed three times in 0.05 M Tris, *p*H 8.3, 0.45 M NaCl and 0.5% NP40, boiled in sample buffer with 1% β-mercaptoethanol and electrophoresed on 5–20% gradient SDS-polyacrylamide gels. The gels were run at 125 V for 1 h after the bromophenol blue marker dye ran off the end of the gel. The gels were then dried, fixed and autoradiographed on Kodak X-R film. Lanes *a*, *c* and are spleen cells; lanes *b*, *d* and *f* are RAW 112 cells; *a* and *b* are precipitated with anti-T200, *c* and *d* with RA3-3A1, and *e* and *f* with control rat IgM. Molecular weights are expressed in thousands (K).

Table 1 Identity of RA3-3A1$^+$ and RA3-2C2$^+$ populations in bone marrow and spleen

Cell fraction	% Cells stained with: RA3-3A1	RA3-2C2	Anti-κ
Whole bone marrow	32	28	10
RA3-3A1$^-$ bone marrow	3	3	1
RA3-2C2$^-$ bone marrow	2	2	1
sIg$^-$ bone marrow	19	19	3
Whole spleen	37	40	45
RA3-3A1$^-$ spleen	1	3	2
RA3-2C2$^-$ spleen	4	3	3
sIg$^-$ spleen	4	5	2

Spleen and bone marrow cells from BALB/c mice were stained with the appropriate antibody followed by a second stage with fluorescein-conjugated rabbit anti-rat Ig (for RA3-3A1 and RA3-2C2) or goat anti-rabbit Ig for the rabbit anti-mouse-κ. The reagents and details of staining and sorting procedures are described elsewhere[11]. The unstained fraction was isolated on the FACS III and samples restained as indicated. Sorted fractions contained 1–3% residual positive cells. The figures represent the per cent positive with the specific antibody minus the per cent positive in a control stain. The control for the rat monoclonals was an unreactive rat IgM monoclonal; for the anti-κ, it was specifically eluted rabbit anti-keyhole-limpet haemocyanin.

largest of which co-migrated with the molecule recognized by RA3-3A1. For RAW 112 cells the anti-T200 precipitated only one wide band centred around MW 220,000, which co-migrated with the band precipitated by RA3-3A1.

To test whether or not the 220,000 MW protein recognized by RA3-3A1 was the same as the large, B cell-associated form of T200, a surface-iodinated spleen cell lysate was precleared with anti-T200 or RA3-3A1 before immunoprecipitation and SDS-PAGE analysis with the two antibodies. Preclearing with anti-T200 removed all the bands precipitable by anti-T200 and also removed all the RA3-3A1-precipitable material (Fig. 2). However, preclearing with RA3-3A1 removed almost all the RA3-3A1-precipitable material, but seemed to remove only part of the 220,000 MW band. This is consistent with the observation (in Fig. 1) that RA3-3A1 does not precipitate as much 220,000 MW material as does anti-T200. This could mean that the 220,000 MW band precipitated by anti-T200 represents two or more antigenically distinct molecules, not all of which are recognized by the RA3-3A1 antibody. Alternatively, the determinant recognized by RA3-3A1 may be more labile than that recognized by the anti-T200, or the RA3-3A1 antibody may have a lower affinity for the molecule than the anti-T200 antibody. The precipitation of comparable amounts of μ chains from untreated and precleared lysates demonstrated that an unrelated surface glycoprotein was not removed by preclearing.

RA3-3A1 was shown to be specific for B cells and B cell precursors by comparing its staining distribution in the spleen and bone marrow with that of an anti-κ antiserum and another monoclonal antibody, RA3-2C2, which has been shown to be specific for B cells in peripheral lymphoid tissues and for B cells

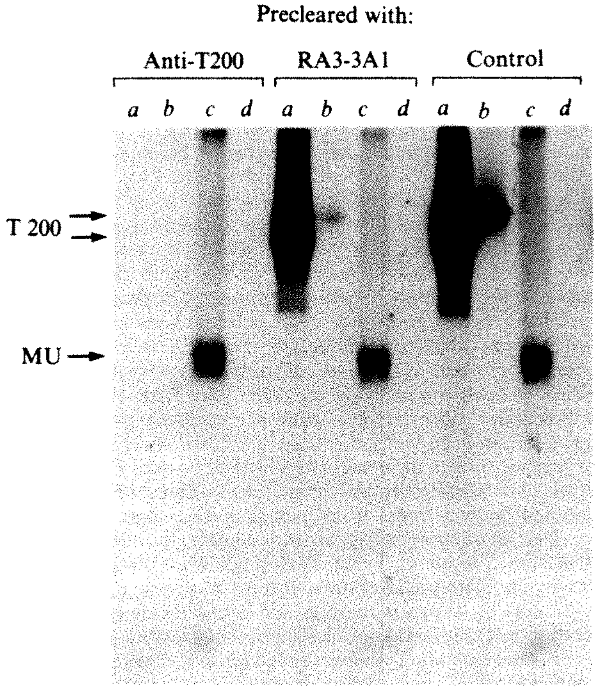


Fig. 2 Preclearing of surface-iodinated spleen lysate with RA3-3A1 and anti-T200. Surface-iodination, antibody precipitation and gel electrophoresis were performed as described in Fig. 1 legend. Preclearing of lysate aliquots was carried out before precipitation by adding a large excess of antibody to the lysates followed by two cycles of precipitation with *S. aureus* (for RA3-3A1 and control IgM) or *S. aureus* saturated with rabbit anti-rat Ig (for anti-T200). Precleared lysates representing 4×10^6 spleen cells were split into four aliquots and immunoprecipitated with: *a*, anti-T200; *b*, RA3-3A1; *c*, monoclonal rat anti-mouse μ chain; and *d*, control rat IgM. The T200 bands were deliberately overexposed to demonstrate the extent of the preclearing. Rat anti-mouse μ chain, produced from the hybridoma 151–119, was given by Dr J. Haaijman.

and sIg$^-$ lymphoid cells in the bone marrow[11]. Removal of sIg$^+$ or RA3-2C2$^+$ cells from the spleen by cell sorting removes all RA3-3A1$^+$ cells and, conversely, removal of the RA3-3A1$^+$ cells removes all cells that are sIg$^+$ and RA3-2C2$^+$ (Table 1). Thus, the molecule recognized by RA3-3A1 is expressed on all splenic B cells. A similar experiment with bone marrow cells again shows that RA3-3A1 and RA3-2C2 recognize the same population of cells, which includes B cells and ~20% of the sIg$^+$ bone marrow cells (Table I). These RA3-2C2$^+$, sIg$^-$ bone marrow cells include those which differentiate into sIg$^+$ cells in two days *in vitro*[11]. RA3-3A1 is therefore expressed on a fraction of sIg$^-$ bone marrow cells which consists of small lymphocytes and medium and large blast cells, and which includes at least some B cell precursors. Despite the similar staining properties of the two monoclonal antibodies, they seem to recognize different molecules. RA3-2C2 precipitates no molecules in the 200,000 MW range from surface-iodinated spleen cells, but it does precipitate a 50,000 MW protein from lysates of biosynthetically labelled lymphoid cells or tumours[12]. This p50 protein is not precipitated by RA3-3A1.

Because the form of T200 recognized by RA3-3A1 is expressed predominantly if not entirely by cells in the B cell lineage, we suggest that this molecule provisionally be called B220, and that the name T200 be reserved for the molecules found on thymus-derived lymphocytes.

We thank Dr William Raschke for the RAW 112 cell line, Drs J. Ledbetter and J. Haaijman for antibodies, and Dr Ian Trowbridge for helpful discussions. This work was supported by NCI fellowships CA 09302 and CA 06429 to R.L.C. and grants CP 91011 from the NCI, AI 09072 from the NIH and IM 56 from the American Cancer Society.

Received 9 October; accepted 2 December 1980.

1. Trowbridge, I. S., Ralph, P. & Bevan, M. J. *Proc. natn. Acad. Sci. U.S.A.* **72,** 157–161 (1975).
2. Trowbridge, I. S. & Mazauskas, C. *Eur. J. Immun.* **6,** 557–562 (1976).
3. Michaelson, J., Scheid, M. & Boyse, E. A. *Immunogenetics* **9,** 193–197 (1979).
4. Omary, M. B., Trowbridge, I. S. & Scheid, M. P. *J. exp. Med.* **151,** 1311–1316 (1980).
5. Siadak, A. W. & Nowinski, R. C. *J. Immun.* **125,** 1400–1401 (1980).
6. Trowbridge, I. S. *J. exp. Med.* **148,** 313–323 (1978).
7. Ralph, P., Nakoinz, I. & Raschke, W. C. *Biochem. biophys. Res. Commun.* **61,** 1268–1275 (1974).
8. Fox, R. I. & Weissman, I. L. *J. Immun.* **122,** 1697–1704 (1979).
9. Kessler, S. W. *J. Immun.* **115,** 1617–1624 (1975).
10. Ledbetter, J. A. & Herzenberg, L. A. *Immun. Rev.* **47,** 63–90 (1979).
11. Coffman, R. L. & Weissman, I. L. *J. exp. Med.* (in the press).
12. Rotter, V., Witte, O. N., Coffman, R. & Baltimore, D. *J. Virol.* **36,** 547–555 (1980).

Induction of cell–cell channel formation by mRNA

G. Dahl, R. Azarnia & R. Werner

Departments of Physiology and Biophysics and of Biochemistry,
University of Miami, School of Medicine, Miami, Florida 33124

Intercellular junctional communication is very common in normal organized tissue[1]. It provides a pathway for transmission of electrical signals, especially in heart muscle[2,3], and may be important in differentiation and growth control[1,4]. The hydrophilic channels which enable cell–cell communication have been well characterized by biophysical methods[5,6], and there is now good evidence that they are contained in the nexus (gap junctions)[3,7–9]. Little, however, is known about the molecular mechanism of biosynthesis of junctional channels. Knowledge in this area has been obtained almost exclusively from experiments with reaggregated cells[10–15], a system complicated by the fact that *de novo* synthesis of channel proteins is obscured by reassembly of pre-existing subunits or utilization of precursors. To avoid these problems, we have now isolated mRNA from cells that are in the process of making new intercellular nexus with high efficiency, incorporated it via liposomes into communication-defective cells and have shown that the recipient cells established junctional communication.

We wanted a system where formation of junctions is controlled by an external stimulus and so chose uterine smooth muscle cells, because their response to stimulus is more pronounced than in the few other cell types where induction has been observed[16–20]. Normally, the cells of the myometrium do not exhibit synchronous contractile activity and have very few gap junctions. By contrast, during parturition when well synchronized activity of the cells of the myometrium is required for expulsion of the fetus, numerous large gap junctions are observed[16,17]. Although we have already reported that nexus formation in the myometrium can be induced in virginal females by injection of oestrogen[16], it was not clear whether this was due to *de novo* synthesis of junctional elements or to reassembly of pre-existing subunits. On the assumption that we were observing *de novo* synthesis of gap junction proteins[21–23] there should be present in these cells an increased number of mRNA molecules coding for these proteins shortly after stimulation with oestrogen.

To investigate this prediction we first obtained either a polysomal fraction or free mRNA from oestrogen-treated uterine tissue of rats. The polysomal fraction, or free mRNA, was then encapsulated into liposomes by the reverse-phase evaporation technique[24]. The phospholipid composition (phosphatidylserine/phosphatidylcholine) of the liposomes was chosen to ensure liposome–cell membrane fusion. Control experiments with liposomes containing fluorescent dyes, and electron microscopy, have shown that fusion was indeed a predominant form of liposome–cell interaction (R.A and G.D., unpublished results) (Fig. 1).

To test whether the injected mRNA coded for gap junction protein we measured its ability to induce intercellular communication between cells of the communication-defective Cl-1D mouse cell line[25]. In all experiments confluent cultures of

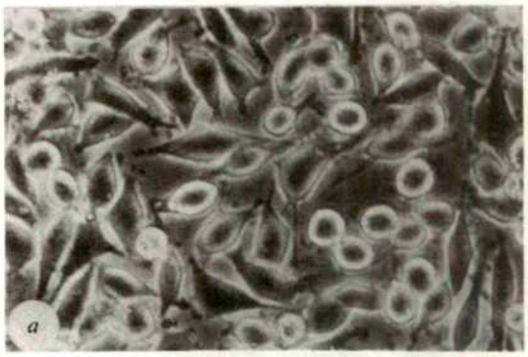


Fig. 1 Appearance of Cl-1D cells in bright field (*a*) and dark field (*b*) after incubation for 10 min with liposomes containing the fluorescent tracer substance carboxyfluorescein. The liposomes were prepared as described for the encapsulation of polysomes except that 100 mM carboxyfluorescein replaced the polysomes. All cells are highly fluorescent and even small cell processes contain sufficient tracer substance to outline them clearly.

0028–0836/81/080683—03$01.00

Cl-1D cells were first washed with serum-free minimal essential medium (MEM). The liposome preparation was diluted 1:3 (v/v) with serum-free MEM, and 1 ml of this suspension was added to a 9.6-cm^2 dish of Cl-1D cells. The incubation period was 1 h at 37 °C. The liposomes were then removed, and the cells washed repeatedly and incubated in normal culture medium. Thereafter random cell pairs were tested for electrical coupling (Fig. 2). Most cell pairs when tested 6 h after the removal of the liposomes were found to be electrically coupled (Table 1). The observed coupling ratios varied, but were consistently <1.0. Control cultures of Cl-1D cells were incubated in identical conditions with liposomes containing buffer only; no coupling was observed. Similarly, if the polysome preparation was treated with ribonuclease before liposomal encapsulation, no coupling developed. When cycloheximide was added to the cells immediately after the 1-h incubation period with mRNA-containing liposomes, the expression of the mRNA was drastically reduced; only 1 out of 13 cell pairs showed any coupling.

Although most of our experiments were done with polysome preparations we have obtained similar results with free mRNA isolated from oestrogen-induced uterine cells (Table 1). In contrast to polysomes, however, free mRNA had to be encapsulated at higher concentrations than polysomes to be as effective in inducing cell–cell coupling. This was not unexpected because the free mRNA was extracted from total cellular RNA whereas polysomes were obtained from a microsomal fraction known to be enriched for mRNA coding for membrane proteins. Thus, the free mRNA may have a lower specific activity than polysomes to induce cell–cell coupling.

Table 1 Incidence of electrical coupling in Cl-1D cells in control conditions and after loading with mRNA

Conditions of liposome treatment	Coupling incidence
Control: liposomes containing buffer only	0/29 (3)
Liposomes containing polysomes	77/114 (14)
Polysomes treated with ribonuclease before encapsulation	0/18 (2)
Cycloheximide (5 μg ml^{-1}) added after polysome loading	1/13 (2)
Free mRNA-containing liposomes:	
5 μg ml^{-1} mRNA	4/20 (2)
25 μg ml^{-1} mRNA	17/22 (2)

The number of cell pairs exhibiting electrical coupling (first number) and the total number of cell pairs impaled with microelectrodes and exhibiting stable membrane potentials (second number) show the coupling incidence. The number of experiments is given in parentheses. Coupling incidence was determined 6 h after treatment of the cells with liposomes. For obtaining the mRNA preparations, mature female virgin rats (Sprague–Dawley) were injected intramuscularly with oestrogen (Delestrogen, Squibb). Each animal received 10 mg of oestrogen per day for 3 days. On the fourth day the rats were killed and the uteri removed. (In some experiments the endometrium was removed by microdissection.) The uterine tissue from five rats was homogenized with an Ultra Turrax (Tekmar) homogenizer in 15–20 ml of buffer A (25 mM sucrose, 1.5 mM $MgCl_2$, 50 mM NaCl, 25 mM HEPES, *p*H 7.2). After centrifugation for 3 min at 5,000*g*, the supernatant was further centrifuged at 100,000*g* for 60 min. The pellet obtained in the high-speed centrifugation (microsomal fraction) was resuspended in 2 ml of buffer B (25 mM NaCl, 5 mM $MgCl_2$, 25 mM HEPES, *p*H 7.2) containing 0.5% (v/v) NP40 detergent. After incubation at 0 °C for 60 min and removal of insoluble fragments by centrifugation (10,000*g*, 30 min), the solution was layered on a discontinuous sucrose gradient containing 1 ml 2.5 M sucrose (in buffer B) and 2 ml 1.0 M sucrose (in buffer B), and centrifuged in a Spinco SW 50.1 rotor at 43,000 r.p.m. for 3 h. A polysomal fraction was recovered from the interface between the two sucrose layers. The polysomes were dialysed overnight against buffer C (1 mM EGTA, 15 mM KCl, 10 mM NaCl, 5 mM HEPES, *p*H 7.2), concentrated by freeze-drying to about 1 ml and dialysed again against buffer C containing 150 mM KCl. The volume of the final polysome preparation was adjusted to a concentration of $A_{260} = 2.0$–3.0 (ratio $A_{260}/A_{280} = 1.5$–1.6). Free mRNA was isolated by homogenization of uterine tissue in 4 M guanidine thiocyanate according to the procedure by Chirgwin *et al.*[27]. All equipment and containers used in these experiments were treated with dietyl pyrocarbonate (0.1%) for inactivation of RNase. The mRNA was further purified by adsorption to oligo(dT) cellulose[28]. For encapsulation of the mRNA preparations into liposomes, the reverse-phase evaporation technique[24] was used with the following specifications: 20 mg of phosphatidylserine and phosphotidylcholine (ratio 3:2) per 0.5 ml of aqueous solution of RNA. The liposome preparation was dialysed against Ca, Mg-free Hank's balanced salt solution (Gibco).

Studies of the time course of development of cell–cell coupling showed that coupling appeared within 2–4 h after removal of the liposomes (Table 2). At 4 h, about 40% of the cells were coupled, and this percentage increased still further until at 8 h practically all cells were coupled. The increase in coupling incidence was accompanied by larger coupling ratios. Later, the frequency of coupling declined and disappeared completely at about 30 h.

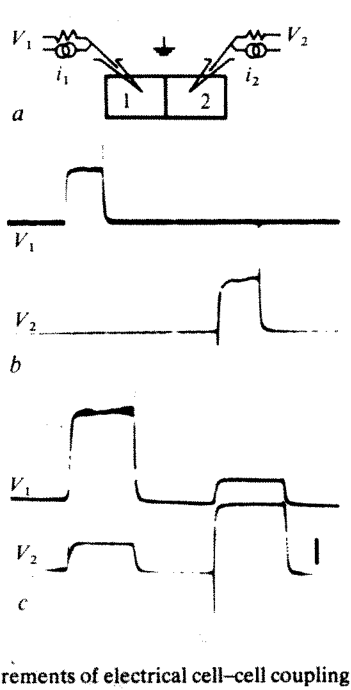


Fig. 2 Measurements of electrical cell–cell coupling. *a*, Electrode arrangement. Two neighbouring cells were impaled with a microelectrode each, which through a bridge circuit allowed current injection and recording of membrane potential. Oscilloscope traces: *b*, membrane potential in Cl-1D cells in control conditions; cells were incubated with liposome-containing buffer solution only. Current injection (10^{-8} A) into cell 1 results in hyperpolarization of this cell (V_1), which does not result in a voltage change in cell 2 (V_2). After a delay of about 150 ms, current is injected into cell 2, which results in change of membrane potential in this cell but not in cell 1. *c*, Membrane potential Cl-1D cells incubated with liposomes containing the mRNA preparation. Current injection in either one of the cells results in potential changes in both cells indicating electrical coupling. Calibration 50 mV.

Freeze-fracture electron microscopy[29] of mRNA-treated Cl-1D cells revealed the presence of nexus-like structures at the time when cell–cell coupling was present (Fig. 3). The observed structures consisted of aggregates of membrane particles resembling nexus which were associated with short ridges appearing like tight junctions. Similar structural arrangements have been found in myometrial tissue during early stages of nexus development after oestrogen treatment (G.D., unpublished results).

To determine whether the observed development of electrical coupling between Cl-1D cells occurs through newly synthesized channel proteins, the following potential artefacts must be ruled out. First, coupling could have developed because of fusion or partial fusion between neighbouring cells. Second, our polysome preparation may have been contaminated by channel

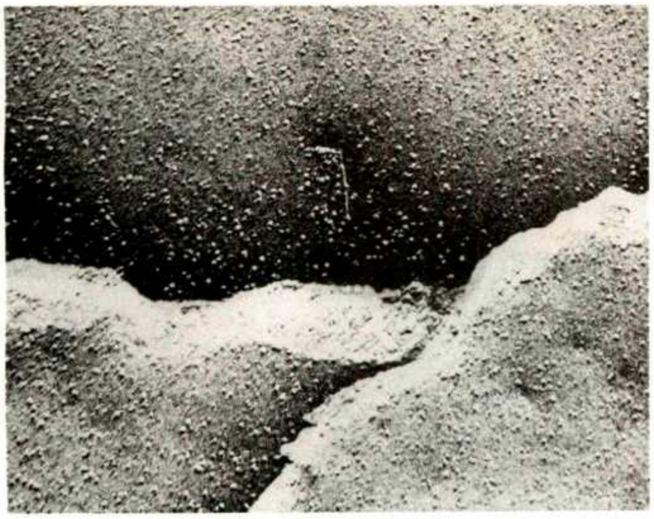


Fig. 3 Freeze-fracture electron micrograph of Cl-1D cells 6 h after incubation with liposomes containing the mRNA preparation. Within an area of membrane apposition a small aggregate of membrane particles can be seen associated with short ridges. ×60,000.

proteins of the uterine cells. Third, the liposome treatment by itself may have nonspecifically induced the development of cell–cell coupling in the recipient cell. Cell–cell fusion can be ruled out, as in all our experiments the observed coupling ratios were considerably below 1.0, whereas fusion should result in coupling ratios of 1.0 or close to 1.0, as has been shown with fusing myoblasts[26]. Furthermore, the control experiments with ribonuclease or cycloheximide indicate that mRNA is required and that protein synthesis is involved in the establishment of coupling. This argument is strengthened by the results obtained with free mRNA. Liposomes by themselves did not produce coupling. The time course of the development of coupling also seems to argue against an artefact. Polysomes from uterine cells that had not been treated with oestrogen did not produce any coupling. Finally, freeze-fracture electron micrographs show membrane junctions. Although the observed structures are not typical for nexus they resemble the structures that can be seen in the uterus shortly after oestrogen treatment. It is possible that the mRNA concentration is too low to induce fully developed nexus. This possibility is supported by the observed coupling ratios which are much lower than those observed in normal, coupled cells. Alternatively, junction formation may be a complex process requiring the participation of several factors some of which may not be under oestrogen control and, consequently, their mRNA may not be present at a high enough concentration in the polysome fraction.

It seems likely, therefore, that functional channels are present in the mRNA-treated Cl-1D cells but are not easily identified by electron microscopy because of the absence of typical multi-unit structures. Assuming that we are observing formation of regular cell–cell channels, we may conclude that their functional lifetime is of the order of 24 h. This seems to be a reasonable estimate because nexus have been observed for a similar period in myometrial cells at term[16]. Of course, the functional lifetime of intercellular channels in organized tissue may be quite different from that observed in cultured cells which are rapidly proliferating and, therefore, have a higher overall turnover rate. In conclusion, our results suggest that the induction of cell–cell coupling in rat uterus smooth muscle by oestrogen occurs by *de novo* synthesis of cell–cell channel proteins. Although this is the most straightforward interpretation of our results, we realize that alternative possibilities cannot be ruled out. For example, the introduced mRNA might code for a regulatory or morphopoietic protein involved in cell–cell communication. We believe that the system described here provides a promising model for obtaining information on the biosynthesis of junctional proteins not readily obtainable by conventional methods.

Table 2 Time course of development of cell–cell coupling after loading with polysomes

Hours after liposome treatment	Coupling incidence
2	0/12
4	4/10
6	9/10
8	10/10
25	3/10
30	0/10

Data obtained from one experiment. Two additional experiments gave similar results.

We thank Dr W. R. Loewenstein for continuous encouragement, and A. Diaz-Calejo and L. Hevia for technical assistance. This research was supported by PHS grants CA 14464, CA 14395 and GM 22510.

Received 21 August; accepted 11 December 1980.

1. Loewenstein, W. R. *Biochim. biophys. Acta* **560,** 1–65 (1979).
2. Weidmann, S. *J. Physiol., Lond.* **118,** 348–360 (1952).
3. Barr, L., Dewey, M. M. & Berger, W. *J. gen. Physiol.* **48,** 797–823 (1965).
4. Loewenstein, W. R. *Devl Biol.* **19,** Suppl. 2, 151–183 (1968).
5. Loewenstein, W. R. *Cold Spring Harb. Symp. quant. Biol.* **40,** 49–63 (1975).
6. Flagg-Newton, J., Simpson, I. & Loewenstein, W. R. *Science* **205,** 404–407 (1979).
7. McNutt, N. S. & Weinstein, R. S. *Prog. Biophys. molec. Biol.* **26,** 45–101 (1973).
8. Caspar, D. L. D., Goodenough, D. A., Malkowsky, L. & Phillips, W. C. *J. Cell Biol.* **74,** 605–628 (1977).
9. Unwin, P. N. T. & Zampighi, G. *Nature* **283,** 545–549 (1980).
10. Loewenstein, W. R. *Devl Biol.* **15,** 503–520 (1967).
11. Ito, S., Sato, E. & Loewenstein, W. R. *J. Membrane Biol.* **19,** 305–337 (1974).
12. DeHaan, R. L. & Hirakow, R. *Expl Cell Res.* **70,** 214–220 (1972).
13. Hammer, M., Epstein, M. & Shridan, J. *J. Cell Biol.* **59,** 130a (1973).
14. Epstein, M. L., Sheridan, J. D. & Johnson, R. G. *Expl Cell Res.* **104,** 25–30 (1977).
15. Griepp, E. B. & Bernfield, M. R. *Circulation* **52,** Suppl. 2, 106 (1975).
16. Dahl, G. & Berger, W. *Cell Biol. int. Rep.* **2,** 381–387 (1978).
17. Garfield, R. E., Sims, S. M., Kannan, M. S. & Daniel, E. E. *Am. J. Physiol.* **235,** C168–C179 (1978).
18. Merk, F. B., Botticelli, C. R. & Albright, J. T. *Endocrinology* **90,** 992–1007 (1972).
19. Decker, R. S. *J. Cell Biol.* **69,** 669–685 (1976).
20. Decker, R. S., Donta, S. T., Larsen, W. J. & Murray, S. A. *J. supramolec. Struct.* **9,** 497–507 (1978).
21. Henderson, D., Eibl, H. & Weber, K. *J. molec. Biol.* **132,** 193–218 (1979).
22. Hertzberg, E. L. & Gilula, N. B. *J. biol. Chem.* **254,** 2138–2147 (1979).
23. Finbow, M., Yancey, S. B., Johnson, R. & Revel, J. P. *Proc. natn. Acad. Sci. U.S.A.* **77,** 970–974 (1980).
24. Szoka, F. & Papahadjopoulos, D. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4194–4198 (1978).
25. Azarnia, R., Larsen, W. J. & Loewenstein, W. R. *Proc natn. Acad. Sci. U.S.A.* **71,** 880–884 (1974).
26. Rash, J. E. & Fambrough, D. *Devl Biol.* **30,** 168–186 (1973).
27. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18,** 5294–5299 (1979).
28. Kryxtosek, A., Cawthon, M. L. & Kabat, D. *J. biol. Chem.* **250,** 6077–6084 (1975).
29. Pauli, B. K., Weinstein, R. S., Soble, L. W. & Alroy, J. *J. Cell Biol.* **72,** 763–769 (1977).

Block of Na current and intramembrane charge movement in myelinated nerve fibres poisoned with a vegetable toxin

J. M. Dubois & M. F. Schneider

Laboratoire de Neurobiologie, Ecole Normale Supérieure,
46, rue d'Ulm, 75230 Paris Cedex 05, France

In nerve membrane, the non-linear capacity current (displacement current) is assumed to reflect the movement of intrinsic membrane charges which control the opening of specific pathways for sodium ions (Na channels). However, various discrepancies have been reported between the effects of pharmacological agents on sodium and displacement currents (for a review see ref. 1). It is generally supposed that the opening and closing of Na channels constitutes one step of a multi-step system in which each configuration change may or not give rise to a measurable charge movement[2]. New drugs affecting either

0028-0836/81/080685—04$01.00

sodium or displacement currents may elucidate the relationship between charge movement and the control of sodium conductance. We therefore now report a comparison of the effects of a vegetable toxin (oenanthotoxin or OETX)[3] on both sodium current (I_{Na}) and intra-membrane charge movement (Q) in Ranvier nodes. We show that OETX reversibly blocks both sodium and displacement current. Studies of I_{Na} and Q during partial suppression by the toxin reveal differences in the effects of OETX on the remaining I_{Na} and Q. The findings are discussed in relation to recent models for the Na-channel gating process and for Na-channel block by local anaesthetics.

Oenanthe crocata is a highly poisonous plant[4,5]. The toxic substance, OETX, isolated from the roots, is a neutral ethylenic and acetylenic diol ($C_{17}H_{22}O_2$). In frog myelinated nerve fibres, OETX (0.1 mM) blocks the action potential and the sodium, potassium and displacement currents (present results and J.M.D., unpublished observations). The present experiments were carried out on voltage-clamped myelinated nerve fibres from the frog *Rana esculenta*. K current was abolished by adding tetraethylammonium (10 mM) to the external solution and by replacing the axoplasmic K content by caesium. Linear capacity and leakage currents were analogically subtracted from the total current[6]. Charge movement was recorded in the presence of tetrodotoxin (TTX, 10^{-6} M). To minimize noise and to obtain direct estimates of charge transfer, the current remaining after subtracting linear components was analogically integrated and then digitized and averaged. Typical recordings of charge transfer (integral of the asymmetrical capacity current) are presented in Fig. 1. In agreement with earlier reports[2,7], charge transfer followed a single exponential time course for moderate depolarizations ($E \leq 0$ mV). For larger depolarizations, two exponentials were required to describe the charge transfer kinetics.

External application of 60 µM OETX reversibly reduced Q (Fig. 1) and I_{Na} (Fig. 2) to 5–10% of their original values. These effects cannot be attributed to either fast (Fig. 3*c*) or slow[8] inactivation because they could not be removed by hyperpolarized holding potentials. The reduction of I_{Na} and Q were both independent of the frequency of stimulation. Recordings of sodium current and charge transfer obtained by intermittent application of saxitoxin during the development of OETX block indicated that OETX decreased I_{Na} and Q in parallel[9]. The leakage and the symmetrical capacity currents remained unchanged when I_{Na} and Q were virtually completely blocked by OETX (60 µM)[9].

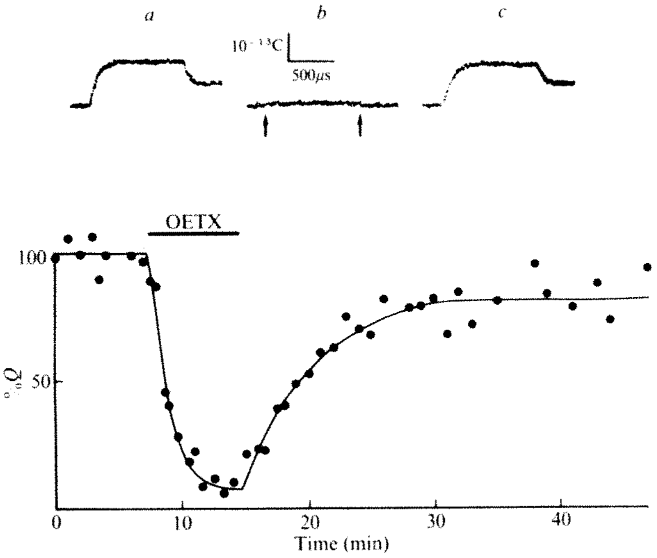


Fig. 1 Reversible immobilization of intra-membrane charges by OETX. The averaged integrated gating current was obtained in response to 12 100-mV depolarizing and 24 50-mV hyperpolarizing pulses from a holding potential of −100 mV. Temperature, 12 °C. Upper: *a*, in the control solution; *b*, in the presence of 60 µM OETX; *c*, after a 24-min wash with the control solution. Lower: relative amplitude of the charge transfer (Q) before, during and after the application of OETX (60 µM). The amplitude of Q was normalized to its mean value before the application of the toxin. Absolute charge and current amplitudes were in all cases calculated assuming a 10-MΩ internal resistance. A 4×10^{-2} solution of OETX in ethanol was diluted in Ringer's solution to give final toxin concentrations. Fibre 6-11-79 A.

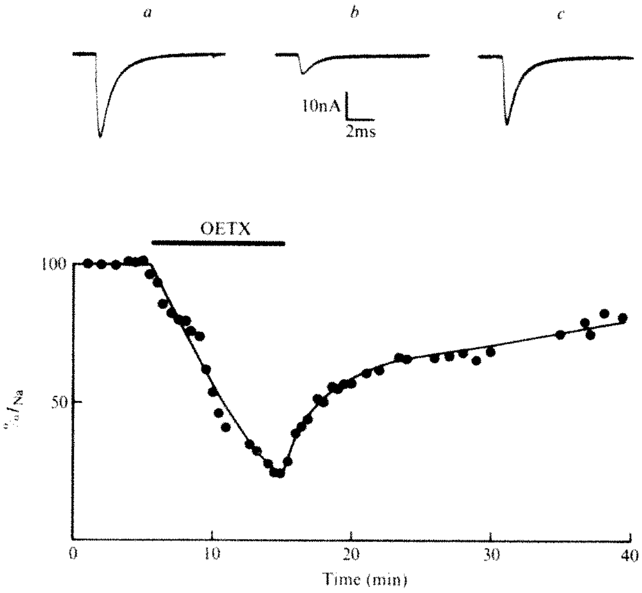


Fig. 2 Reversible block of I_{Na} by OETX. Sodium current was recorded at $E = 0$ mV after a 50-ms prepulse to −110 mV (control) or −140 mV (OETX). Holding potential, −70 mV. Temperature, 12 °C. Upper: *a*, in the control solution; *b*, in the presence of 60 µM OETX; *c*, after a 23-min wash with control solution. Lower: relative amplitude of peak I_{Na} before, during and after the application of OETX (60 µM). The amplitude of I_{Na} was normalized to its mean value before the application of the toxin. Fibre 13-11-79.

In addition to reducing I_{Na} and Q, OETX altered various properties of non-blocked Na channels. The fractional depression of peak I_{Na} produced by OETX varied with voltage (Fig. 3*a*), indicating that the Na-permeability activation curve was decreased in slope with a shift of its 50% point towards positive voltages (Fig. 3*b*). Similar results were obtained in four other fibres in which this point was investigated. In contrast, the position and relative steepness of the Q_{on}-voltage curve were not significantly modified by toxin treatment which reduced Q_{on} at 0 mV to 34–57% of control in three fibres (Fig. 4*a*). This is shown by Fig. 4*b*, in which the Q_{on} results in OETX are replotted with values for each fibre scaled up so as to compensate for the depression of Q by OETX. Both the data points in Fig. 4*b* and their best fit (dashed curve) by the two-state Boltzmann model (Fig. 4 legend) lie close to the fit (solid curve) to the control Q_{on} data from the same fibres which are reproduced from Fig. 4*a* in 4*b*. The effect of OETX on the Na-permeability activation curve cannot be attributed to a decrease in the series resistance artefact due to the smaller I_{Na} (ref. 10) because a reduction of I_{Na} to 40% of its original value by TTX (4×10^{-9} M) caused the Na-permeability activation curve to be shifted towards positive voltages by less than 2 mV (refs 9, 11). The time course of Na-current activation, determined after correction for a two-exponential inactivation process[12] in parallel, was either slightly delayed or not significantly affected by OETX[9], whereas the time constants for charge movement were reduced[9].

OETX markedly shifted the Na-inactivation curve towards negative potentials and slightly decreased its slope (Fig. 3*c*). The

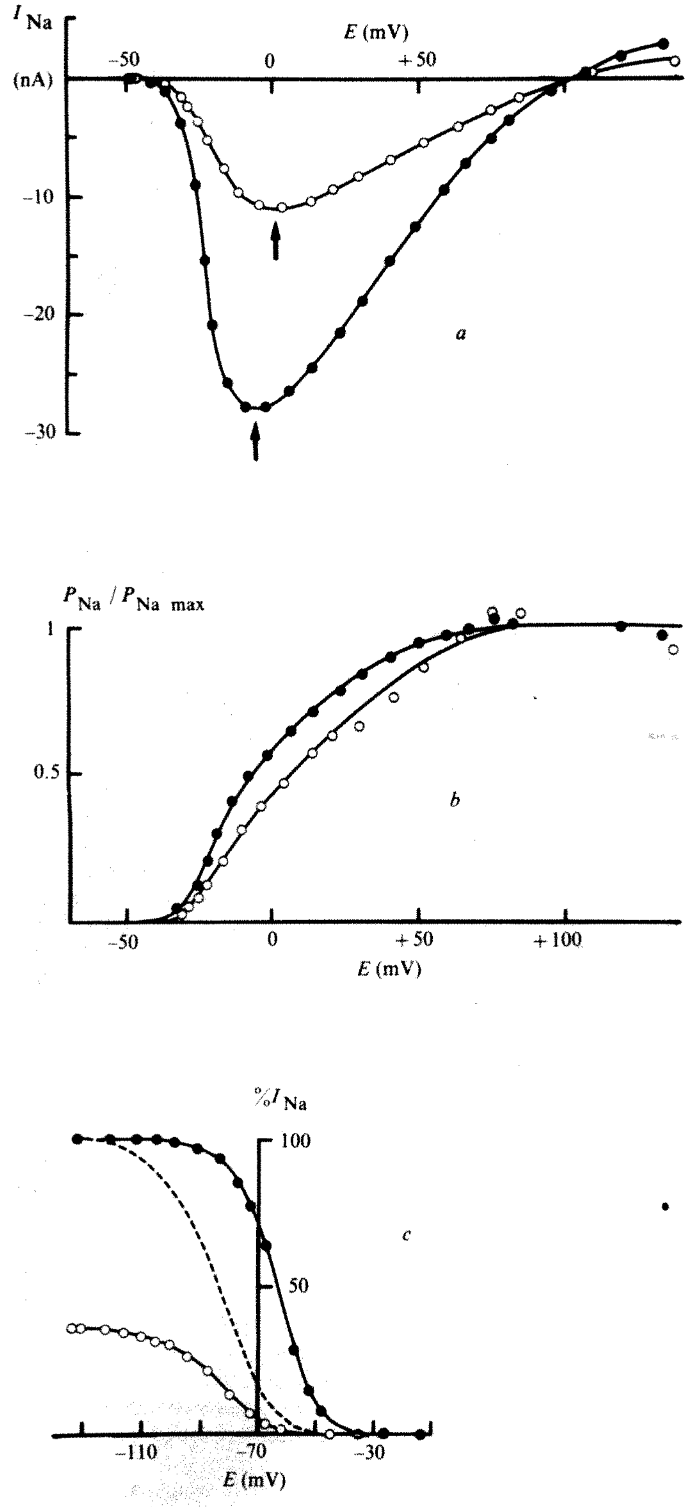


Fig. 3 Na current (*a*), Na permeability (*b*) and Na-current inactivation (*c*) voltage curves in control solution (filled circles) and during wash shortly after exposure to OETX (200 μM) (open circles). Holding potential, −70 mV. Temperature, 12 °C. *a*, The peak of I_{Na} was recorded during test pulses preceded by 50-ms prepulses to −110 mV (control solution) or −140 mV (during wash), so as to ensure complete removal of Na inactivation. Note that under the effect of toxin, the potential at which the inward current is maximum (arrows) is shifted towards positive values. *b*, Relative peak Na permeability (P_{Na}) as a function of membrane potential calculated from each set of results in *a* assuming the Na channel to rectify according to the constant field equation[19]. *c*, Available Na current (per cent of maximum control I_{Na}) as a function of membrane potential during a 50-ms prepulse. The dashed curve gives the results in OETX normalized to a maximum value of 100% for large hyperpolarizations. I_{Na} was measured at 0 mV. Fibre 26-2-79.

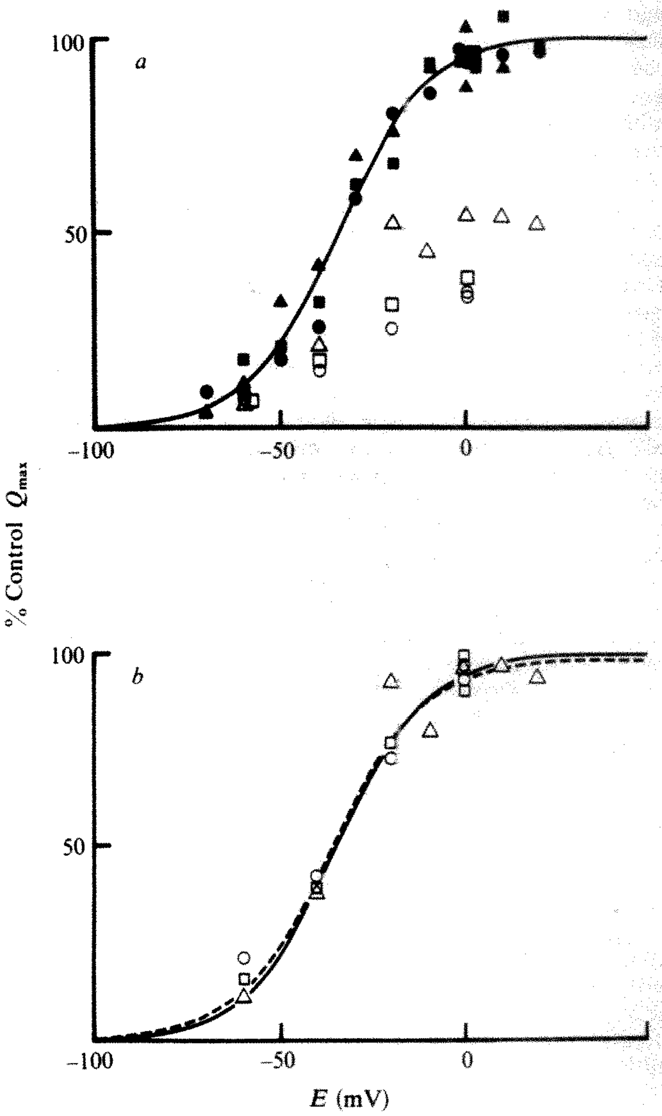


Fig. 4 Relative charge–voltage relationship in control solution (filled symbols) and in the presence of 40–60 μM OETX (open symbols). Holding potential, −100 mV. Temperature, 10–12 °C. *a*, Values obtained in both solutions on each of three fibres (different symbols) were first normalized to the mean control value of Q_{on} in that fibre at 0 mV. A non-linear least squares fit of the equation

$$Q_{on} = Q_{max}/\{1+\exp[(\bar{E}-E)/k]\} \qquad (1)$$

for the two-state Boltzman model[20–22] was first made to the combined values from the three fibres in control solution. All Q_{on} values were then normalized to the control Q_{max}. The values of Q_{max} (normalized), $\bar{E}$ and k obtained from the fit (curve) were 100%, −34.3 mV and 12.0 mV, respectively. *b*, Q values in OETX for each fibre were scaled by a constant chosen so as to minimize the squared differences between scaled values and the fit to the combined control data in *a*. The fractional values of Q in OETX determined from these scale factors were 35% (○), 57% (△) and 40% (□) of the control fit. Equation (1) was then fitted to the combined scaled OETX results, giving the dashed curve (Q_{max} = 98.1%, $\bar{E} = -36.4$ mV and $k = 12.6$ mV). The fit to the control data (solid curve) is reproduced from *a* for comparison. Fibres 15-10-79 (○), 6-11-79A (△), 21-12-79 (□).

kinetics of Na inactivation were accelerated[9] and the ratio Q_{off}/Q_{on} for a given pulse duration was slightly decreased.

Two major conclusions can be drawn from these results: OETX induces a reversible block of I_{Na} and charge movement, and in the presence of the toxin, the voltage and time dependences of the remaining sodium current and charge movement seem to be altered differently. Because the latter conclusion is based on measurements of OETX effects in different fibres with probable differences in toxin exposure, a quantitative comparison of toxin effects on I_{Na} and Q was not attempted.

In some respects, the effects of the uncharged toxin OETX seem to resemble those of the uncharged local anaesthetic benzocaine, which depresses both I_{Na} (refs 8, 13) and Q (ref. 14) and shifts and decreases the slope of the Na-inactivation curve[13]. However, whereas maximum attainable concentrations of benzocaine produce only partial depression of sodium current[8,15,16] and charge movement (ref. 14 and our unpublished observations), OETX can almost completely suppress I_{Na} and Q. This difference may be due to an inability to achieve sufficiently high benzocaine levels at the reactive channel site for all channels to react with drug, whereas OETX is sufficiently more soluble in the membrane or more reactive with the site to occupy almost all the sites.

In view of the similarities between the effects of OETX and benzocaine, it seems likely that the OETX effects on steady-state parameters may be accounted for by a general scheme for local anaesthetic action in which the drug can bind to the Na channel in each of its several states[13]. The depression of I_{Na} which cannot be restored by hyperpolarization would be due to OETX binding to resting channels. The shift and change in slope of the Na-inactivation curve would be explained by drug binding to open and inactivated channels. With this interpretation, the depression of Q would indicate that drug-blocked channels cannot undergo transitions between states, at least not transitions giving rise to detectable charge movement. Such a scheme for OETX action may perhaps also account for the modification of kinetic parameters. Alternatively, in addition to binding to the channel in its various states, it is possible that OETX produces the observed kinetic effects by acting at a different site from that at which it modifies the rate constants for channel transitions. In either case, the possibility of multiple transitions between resting and activated states[2,17], which may contribute differing amounts of Q, may perhaps account for the differences observed between OETX effects on the kinetics of I_{Na} and Q and on the voltage dependencies of sodium permeability and Q. The accelerated Na inactivation and decreased Q_{off}/Q_{on} are consistent with the suggested relationship between inactivation and charge immobilization[2,18].

We thank J. Louvel and Y. Chauvel for providing OETX samples and N. L. Sizto for running the curve fitting program. M.F.S. was on leave from the Department of Physiology, University of Rochester, Rochester, New York 14642.

Received 23 June; accepted 2 December 1980.

1. Almers, W. *Rev. Physiol. Biochem. Pharmac.* **82,** 96–190 (1978).
2. Armstrong, C. M. & Bezanilla, F. *J. gen. Physiol.* **70,** 567–590 (1977).
3. Anger, J. P. *et al. Eur. J. Tox.* **9,** 119–125 (1976).
4. Grundy, H. F. & Howarth, F. *Br. J. Pharmac.* **11,** 225–230 (1956).
5. Chauvel, P., Louvel, J., Anger, J. P. & Chauvel, Y. *C. r. hebd. Acad. Sci. Séance, Paris* **286,** 1833–1836 (1978).
6. Dubois, J. M. & Bergman, C. *Pflügers Arch. ges. Physiol.* **370,** 185–194 (1977).
7. Nonner, W., Rojas, E. & Stämpfli, R. *Pflügers Arch. ges. Physiol.* **375,** 75–85 (1978).
8. Khodorov, B., Shishkova, L., Peganov, E. & Revenko, S. *Biochim. biophys. Acta* **433,** 409–435 (1976).
9. Dubois, J. M. & Schneider, M. F. *Adv. physiol. Sci.* (in the press).
10. Drouin, H. & Neumcke, B. *Pflügers Arch. ges. Physiol.* **351,** 207–229 (1974).
11. Neumcke, B., Nonner, W. & Stämpfli, R. *Pflügers Arch. ges. Physiol.* **363,** 193–203 (1976).
12. Chiu, S. Y. *J. Physiol., Lond.* **273,** 573–596 (1977).
13. Hille, B. *J. gen. Physiol.* **69,** 497–515 (1977).
14. Khodorov, B., Guselnikova, G. & Peganov, E. *Dokl. Acad. Nauk SSSR* **244,** 1252–1255 (1979).
15. Arhem, P. & Frankenhaeuser, B. *Acta physiol. scand.* **91,** 11–21 (1974).
16. Hille, B. *J. gen. Physiol.* **69,** 475–496 (1977).
17. Armstrong, C. M. & Gilly, W. F. *J. gen. Physiol.* **74,** 691–711 (1979).
18. Nonner, W. *J. Physiol., Lond.* **299,** 573–603 (1980).
19. Dodge, F. A. & Frankenhaeuser, B. *J. Physiol., Lond.* **148,** 188–200 (1959).
20. Schneider, M. F. & Chandler, W. K. *Nature* **242,** 244–246 (1973).
21. Keynes, R. D. & Rojas, E. *J. Physiol., Lond.* **239,** 393–434 (1974).
22. Nonner, W., Rojas, E. & Stämpfli, R. *Pflügers Arch. ges. Physiol.* **354,** 1–18 (1975).

Fibrinogen is the receptor for the endogenous lectin of human platelets

T. Kent Gartner*, J. M. Gerrard†, J. G. White & D. C. Williams*

* Department of Biology, Memphis State University, Memphis, Tennessee 38152
† Department of Pediatrics, University of Minnesota Health Science Center, Minneapolis, Minnesota 55455

Washed platelets activated by α-thrombin, γ-thrombin[1], thrombocytin[2] or the ionophore A23187 (ref. 3) lose their disk shape[4], produce pseudopodia and become cohesive. This cohesiveness is accompanied by the expression of an endogeneous haemagglutinin[5] which, although apparently bound to the platelet membrane[6], is dependent on cell secretion[7]. The interaction of this agglutinin with appropriate receptors on other platelets is believed to be responsible for aggregation. We report here that platelets can be prepared which lack agglutinin activity but have receptor function, that afibrinogenaemic platelets lack receptor activity, and that fibrinogen is the receptor for the agglutinin secreted by activated platelets.

Washed, fixed and non-fixed platelets were examined for thrombin-enhanced agglutinin activity using microtitre plates (Table 1). It is known that non-activated platelets do not express the agglutinin[5], and accordingly non-activated washed platelets did not agglutinate either themselves or fixed platelets. In contrast, platelets activated by thrombin agglutinated themselves and their agglutination titre was increased two- to four-fold by the presence of fixed platelets. The formaldehyde-fixed platelets, when treated with thrombin, neither agglutinated themselves nor fixed bovine erythrocytes which had receptors for the platelet agglutinin. Thus, the formaldehyde-fixed platelets had no agglutinin activity, which means that the apparent increase in the agglutination titre resulting from the presence of the fixed platelets was due to agglutination of the latter by the agglutinin of the thrombin-activated washed platelets. These results demonstrate that fixed platelets have functional receptors for the agglutinin but no functional agglutinin and that they did not cause activation of the washed platelets.

Agglutination of the fixed platelets by the activated washed platelets in the microtitre plate assays was blocked by the same compounds which block thrombin-induced platelet aggregation and enhanced haemagglutination[5] (see Table 2). Amino sugars and arginine completely abolished this activity whereas neutral sugars and other amino acids had no effect. Therefore, the binding specificity of the receptor on fixed platelets seems to be indistinguishable from that of either washed platelets or platelets in plasma[5].

In the course of characterizing platelets from patients with bleeding disorders, it was discovered that fixed platelets from a patient with congenital afibrinogenaemia[8] failed to enhance the agglutination titre of thrombin-activated platelets, which means that the fixed afibrinogenaemic platelets lacked agglutinin receptor function (Table 3). Furthermore, it was observed that this deficiency could be corrected if the afibrinogenaemic platelet-rich plasma was diluted 1:1, before fixation, with platelet-poor plasma from a donor who had a normal level of plasma fibrinogen. These results, and evidence that platelets from the afibrinogenaemic patient have no obvious plasma membrane deficiencies (D. R. Phillips, personal communication), led us to speculate that fibrinogen is the receptor for the secreted agglutinin expressed by activated platelets.

This hypothesis was tested by determining if fibrinogen could inhibit the expression of the enhanced haemagglutinin activity of activated platelets, the rationale being that free fibrinogen

0028-0836/81/080688—03$01.00

Table 1 Fixed platelets lack agglutinin activity in microtitre plate assays

Diluted material	Fixed cells added	Titre
Washed platelets	None	0
Thrombin-activated washed platelets	None	2
Thrombin-activated washed platelets	Bovine erythrocytes	16*
Fixed platelets	None	0
Fixed platelets plus thrombin	None	0
Fixed platelets plus thrombin	Bovine erythrocytes	0

Agglutination of platelets was assayed in microtitre 'V' plates (Cooke Engineering Laboratory Products)[6]. 25 μl of Tris-buffered saline containing calcium (TCS buffer: 10 mM Tris, 2 mM $CaCl_2$, 150 mM NaCl, pH 7.4) was added to each well then 25 μl of the diluted material was added to the first in a row of eight wells and was serially diluted (× 2) in a series of six wells. The remaining two wells contained no diluted materials and served as controls. Finally, 25 μl of the fixed cells added were placed into each well, as either a suspension of fixed platelets (1×10^8 ml^{-1}) or a 2.5% suspension of trypsin-treated, formalinized bovine erythrocytes in TCS buffer[6]. Wells not containing fixed cells received 25 μl of TCS buffer and the contents of the wells were agitated gently. Each test was run in triplicate. Well contents were evaluated for agglutination after standing for 7 h at room temperature. The titre of a diluted material is defined as the reciprocal of the end point dilution (greatest dilution that gives agglutination). The platelets used as diluted materials were suspended in Tris-buffered saline containing EDTA (ETS buffer: 1 mM EDTA, 10 mM Tris, 150 mM NaCl, 0.2% α-D-glucose, pH 7.4) at a concentration of 10^9 ml^{-1}. Washed and fixed platelets were treated with thrombin by adding a final concentration of 0.25 U ml^{-1} of human thrombin (provided by Dr J. Fenton II) to the platelets, incubating the mixture for 1 min at 22 °C, adding it to the well and diluting immediately. Platelets were fixed by making a 1 : 1 dilution of platelet-rich plasma, obtained from blood drawn into ACD (2.5 g Na_3 Citrate. $2H_2O$, 2 g glucose, 1.5 g citric acid in 100 ml H_2O), into modified phosphate-buffered saline (40 mM PO_4, 150 mM NaCl, pH adjusted to 7.4 with NaOH) containing 0.5% formaldehyde and incubating overnight with gentle shaking. All operations were done at 37 °C. The fixed platelets were washed three times with TS buffer (10 mM Tris, 150 mM NaCl, pH 7.4) and suspended in TCS before use. Fixed platelets were used on the same day as they were washed. Suspensions of fixed non-activated platelets were birefringent, and the platelets had a disk shape when viewed in a scanning electron microscope. ADP treatment of the platelets (final concentration 5 μM ADP) before fixation typically decreased their receptor function—this may be caused by the generation of excess receptor activity but has not been tested.

* This agglutination was inhibited by a 30 mM final concentration of each of the following: L-arginine, α-D-mannosamine, α-D-glucosamine, and α-D-galactosamine but not by a final concentration of 60 mM glutamine, N-acetylglucosamine or N-acetylgalactosamine.

Table 2 Fixed platelets have receptor activity for the thrombin-enhanced agglutinin

Diluted material	Fixed cells added	Titre
Washed platelets	None	0
Washed platelets	Platelets	0
Thrombin-activated washed platelets	None	2
Thrombin-activated washed platelets	Platelets	8*

Assays were performed as described in Table 1 legend.

* This agglutination was inhibited by the same compounds which inhibit thrombin-induced platelet aggregation and the function of the platelet-mediated enhanced haemagglutination activity of thrombin-activated washed platelets[5], that is, by a 30 mM final concentration of each of the following: L-arginine, α-D-mannosamine, α-D-glucosamine, and α-D-galactosamine but not by a final concentration of 69 mM glutamine, N-acetylglucosamine or N-acetylgalactosamine.

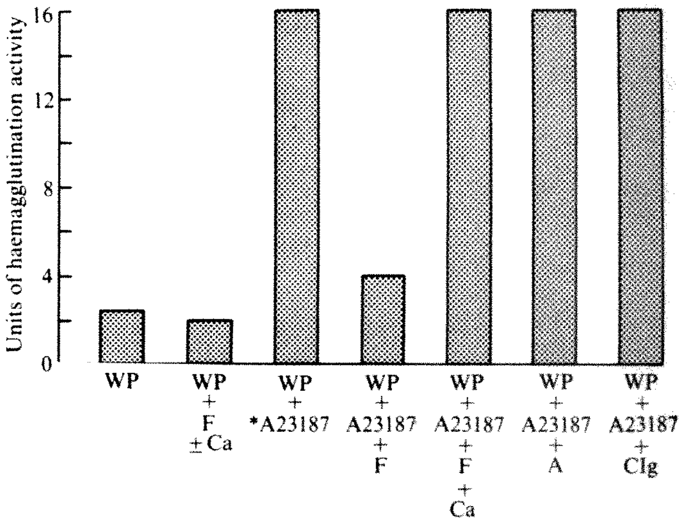


Fig. 1 Microtitre plate assays were carried out as described in Table 1. TS was the plate buffer and calcium ($CaCl_2$) was added where indicated at a final concentration of 2 mM. Platelets suspended in ETS were activated with one of the following: A23187 at 1 μg ml^{-1} (Calbiochem), α-thrombin at 0.25 U ml^{-1}, γ-thrombin at 5 μg ml^{-1} (both provided by Dr J. Fenton II), and thrombocytin at 5 μg ml^{-1} (supplied by Dr K. Stocker, Pentapharm). Human fibrinogen (L-grade, Kabi) treated with diisopropyl fluorophosphate[13], was used at a final concentration of 250 μg ml^{-1}. 125 μg ml¯1 of fibrinogen caused inhibition of half the maximum haemagglutination activity. Similar results were obtained using washed platelets activated with each stimulus listed above. Cold insoluble globulin (CIg, supplied by Dr K. Yamada[14]) was used at a final concentration of 125 μg ml^{-1}. Human serum albumin (Sigma) was used at a final concentration of 250 μg ml^{-1}. WP, washed platelets; F, fibrinogen; Ca, calcium; A, albumin; CIg, cold insoluble globulin. * The haemagglutination caused by washed platelets activated by A23187 was not inhibited by the supernatant fraction from heat-precipitated fibrinogen.

could compete with cells for the secreted agglutinin, if it functions as an agglutinin receptor. Washed platelets activated with α-thrombin, γ-thrombin, thrombocytin or A23187 were tested for enchanced haemagglutination activity in the presence and absence of fibrinogen. As shown in Fig. 1, fibrinogen completely inhibited the expression of the agglutinin in the absence of Ca^{2+} but had no effect in the presence of Ca^{2+}. The fact that fibrinogen did not inhibit the function of the agglutinin in the presence of Ca^{2+} explains why plasma fibrinogen would not be expected to inhibit the function of the endogenous platelet agglutinin *in vivo*. The inhibitory effect seems to be specific for fibrinogen because neither albumin nor cold insoluble globulin caused inhibition. These results may partially explain the observation that monovalent anti-fibrinogen antibody fragments drastically inhibited the aggregation of thrombin-activated washed platelets[9]. The basis of the ability of Ca^{2+} to eliminate the fibrinogen-induced inhibition of enhanced haemagglutination activity is unknown, but a possible explanation is that fibrinogen bound to platelets in the presence of Ca^{2+} has a much greater affinity for the agglutinin than does free fibrinogen.

The possibility that fibrinogen is the haemagglutinin was evaluated by the following tests. Fibrinogen itself was found to have no haemagglutination activity[7] (Fig. 1). Washed platelets suspended in ETS (see Table 1 legend) or Tyrode buffer were treated with fibrinogen and ADP (platelets in Tyrode buffer aggregated in response to stirring in the test conditions) and found to have no enhanced haemagglutination activity. Thus there is no evidence that plasma fibrinogen is the agglutinin, but our results do not exclude the possibility that platelet fibrinogen is the agglutinin. However, as there is evidence that platelet and plasma fibrinogen are identical[10], this possibility will not be considered further. The observations that amino sugars inhibit the binding of fibrinogen to ADP-treated platelets[11] does not elucidate whether fibrinogen functions as an agglutinin rather than a receptor for the agglutinin, but it does explain how these compounds can inhibit platelet aggregation[5].

Table 3 Agglutinin receptor activity of formaldehyde-fixed afibrinogenaemic platelets

Diluted material	Fixed cells added	Titre
Washed platelets	None	0
Thrombin-activated washed platelets	None	2
Thrombin-activated washed platelets	Control platelets	8
Thrombin-activated washed platelets	Afibrinogenaemic platelets	2
Thrombin-activated washed platelets	Afribinogenaemic platelets fixed in diluted control platelet-poor plasma	8

Assays were performed as described in Table 1 legend. The patient with congenital afibrinogenaemia had <10 μg ml^{-1} of plasma fibrinogen and <0.6 μg fibrinogen per 10^8 platelets.

The results described here demonstrate that the agglutinin of activated platelets is distinct from its receptor and that fibrinogen is the receptor for the agglutinin. According to these results and those already published[5,7], secreted materials, including proteins, mediate platelet aggregation by cross-linking adjacent platelets. This interaction of secreted agglutinin with platelets is mediated by fibrinogen—that is, fibrinogen links the agglutinin to the platelet plasma membranes. Fibrinogen, in turn, is probably specifically bound to the plasma membranes via membrane glycoproteins. The identity of these glycoproteins has not been reported, but evidence suggests that platelet glycoproteins IIb and/or III may be the fibrinogen receptors[12].

This work was supported by USPHS grants HL 23010 and HL 11880.

Received 30 June; accepted 9 December 1980.

1. Fenton, J. W. II, Landis, B. H., Walz, D. A. & Finlayson, J. S. in *Chemistry and Biology of Thrombin* (eds Lundblad, R. L., Fenton, J. W. II & Mann, K. G.) 43–70 (Ann Arbor Science, Michigan, 1977).
2. Kirby, E. P. *et al. Biochemistry* **18,** 3564–3570 (1979).
3. Feinman, R. D. & Detwiler, T. C. *Nature* **249,** 172–173 (1974).
4. White, J. G. *Am. J. Path.* **71,** 363–378 (1979).
5. Gartner, T. K., Williams, D. C., Minion, F. C. & Phillips, D. R. *Science* **200,** 1281–1283 (1978).
6. Gartner, T. K., Williams, D. C. & Phillips, D. R. *Biochem. biophys. Res. Commun.* **79,** 592–599 (1977).
7. Gartner, T. K., Phillips, D. R. & Williams, D. C. *FEBS Lett.* **113,** 196–200 (1980).
8. Shapiro, R. S. *et al. Am. J. Pediat. Hemat. Oncol.* **2,** 207–212 (1980).
9. Tollefsen, D. M. & Majerus, P. W. *J. Clin. Invest.* **55,** 1259–1268 (1975).
10. Doolittle, R. F., Takagi, T. & Cottrell, B. A. *Science* **185,** 368–370 (1974).
11. Kinlough-Rathbone, R. L., Chahil, A., Perry, D. W., Packham, M. A. & Mustard, J. F. *Blood* **54,** 249 (1979).
12. Bennett, J. S. & Vilaire, G. *J. clin. Invest.* **64,** 1393–1401 (1979).
13. Mustard, J. F., Perry, W., Kinlough-Rathbone, R. L & Packham, M. A. *Am. J. Physiol.* **228,** 1757–1765 (1975).
14. Yamada, K. M. & Kennedy, D. W. *J. Cell Biol.* **80,** 492–498 (1979).

Non-uniform ion distributions and electrical potentials in sarcoplasmic regions of skeletal muscle fibres

D. G. Stephenson, I. R. Wendt & Q. G. Forrest

Department of Zoology, La Trobe University, Bundoora 3083, Australia

We report here electrophysiological and ion distribution results obtained with mechanically skinned skeletal muscle fibres which indicate that a Donnan equilibrium, characterized by large ionic concentration and electrical potential differences, is established between the myofibrillar space and the rest of the sarcoplasm surrounding the myofibrils. The more negative electrical potential within the myofibrils (−15 to −20 mV for conditions similar to those *in vivo*) can strongly influence: (1) intracellular Ca^{2+} movements associated with contraction; (2) the location of sarcoplasmic enzymes, their substrates and products of reaction; (3) measurements of the electrical potential difference between the extra- and intracellular space and estimation of relative membrane permeabilities for certain ions.

We used mechanically skinned single muscle fibres of the barnacle[1] (*Balanus nigrescens*), subsequently treated with a nonionic detergent which disrupted and removed all cellular membranes. Electron microscopic observations confirmed this and showed that the only regions which could be distinguished in the cross-section of the preparations after this procedure were the myofibrils and the inter-myofibrillar space.

Potential differences were measured between two glass micro-electrodes impaled into the skinned muscle fibre. As can be seen in Fig. 1, one could record either a very small or a relatively large potential difference regardless of whether both electrodes were in the preparation (Fig. 1*A, D*), or only one electrode was impaled in the skinned fibre while the reference electrode was kept in the surrounding solution (Fig. 1*B*). The results indicate that one of the regions within these skinned muscle fibres had essentially the same electrical potential as the surrounding solution, but that a second region yielded negative potentials of large amplitude relative to the bathing solution. When using low-resistance microelectrodes with an external tip similar in size to the diameter of myofibrils (2–3 μm), we were unable to record the large potential differences. This suggests that the former region may be associated with the space around myofibrils, whereas the latter may represent the myofibrils themselves.

The potential differences measured in these skinned muscle fibres did not change following the partial replacement of K^+

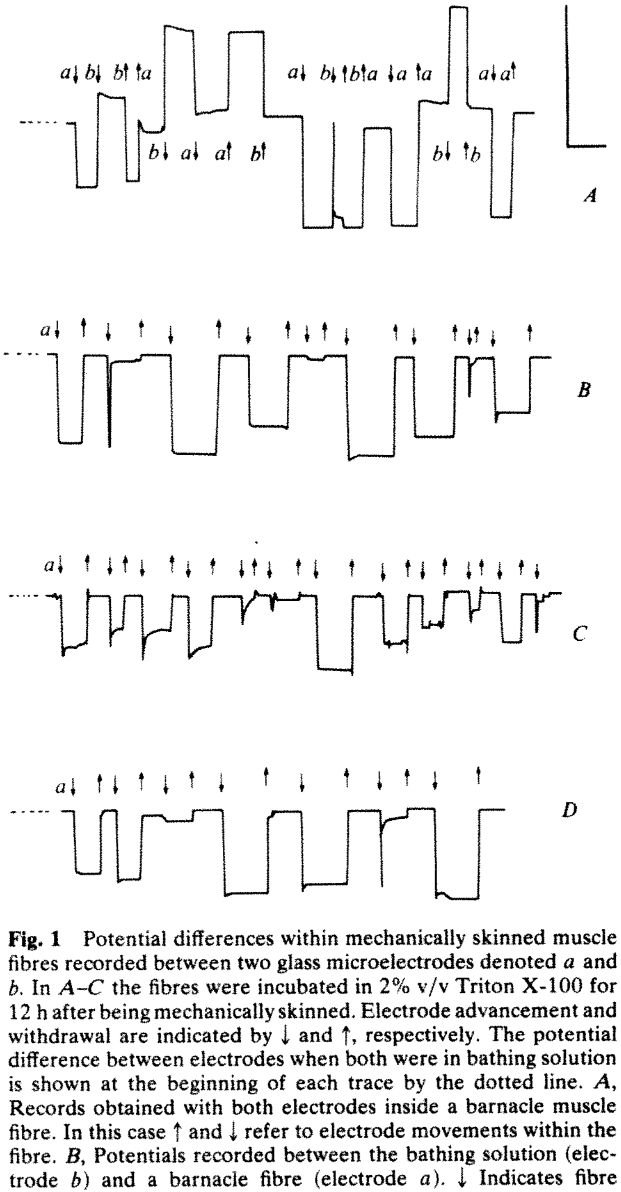


Fig. 1 Potential differences within mechanically skinned muscle fibres recorded between two glass microelectrodes denoted *a* and *b*. In *A–C* the fibres were incubated in 2% v/v Triton X-100 for 12 h after being mechanically skinned. Electrode advancement and withdrawal are indicated by ↓ and ↑, respectively. The potential difference between electrodes when both were in bathing solution is shown at the beginning of each trace by the dotted line. *A*, Records obtained with both electrodes inside a barnacle muscle fibre. In this case ↑ and ↓ refer to electrode movements within the fibre. *B*, Potentials recorded between the bathing solution (electrode *b*) and a barnacle fibre (electrode *a*). ↓ Indicates fibre impalement and ↑ withdrawal of the electrode into the bathing solution. *C*, Potential differences recorded from a toad iliofibularis fibre as in *B*. *D*, Records from a freshly skinned barnacle muscle fibre immersed in oil. In this instance the dotted line represents the potential difference between electrodes when just in contact with the fibre surface. The internal resistance of the glass microelectrodes (normally 40–80 MΩ when filled with either 2 M KCl or 4 M Na-acetate) was routinely measured when the electrodes were in solution and in the preparation to detect and discard artefactual potential measurements due to electrode tip blockage and/or damage. The external tip diameters and tip potentials of these electrodes were characteristically 0.1–0.3 μm and <5 mV, respectively. Input impedance $>10^{13}$ Ω. Composition of bathing solution for *A–C* (mM): EGTA, 10; Na_2ATP, 3; $MgCl_2$, 2; K^+, 100; trimethylaminoethanesulphonic acid (TES), 40; *p*H 7.10 adjusted with Cl. Calibration bars: vertical 30 mV; horizontal 30 s.

Table 1 R_i values for different substances obtained in skinned muscle preparations from nine barnacles in different ionic conditions and predicted potential differences ($\Delta V^i_{\mathrm{pred}}$) between the myofibrillar and extra-myofibrillar space

Γ/2 (mM)*	R_K	R_{EGTA}	R_{Gly}	R_{BSA}	R_{Carb}	R_{Mk}	R_{Myo}	R_{Chym}	$R_{Cyt\,c}$	R_{Aldo}	$\Delta V^K_{\mathrm{pred}}$ (mV)†	$\Delta V^{EGTA}_{\mathrm{pred}}$ (mV)†
55	1.62 ± 0.03	0.53 ± 0.02	0.99 ± 0.01	0.53 ± 0.02	0.73 ± 0.04	—	1.2 ± 0.03	—	4.6 ± 0.2	1.27 ± 0.03	−22 ± 2	< −30
80	1.45 ± 0.02	0.59 ± 0.02	0.98‡	0.57 ± 0.02	—	0.77 ± 0.02	—	—	—	—	−19 ± 2	< −25
120	1.30 ± 0.02	0.69 ± 0.03	0.97 ± 0.02	0.60 ± 0.01	0.79 ± 0.02	—	0.98 ± 0.03	2.06 ± 0.02	2.4 ± 0.1	0.85 ± 0.02	−16 ± 2	−18 ± 5
220	1.17 ± 0.02	0.77 ± 0.02	0.966‡	0.63 ± 0.01	—	0.82 ± 0.02	—	—	—	—	−12 ± 1.5	−11 ± 3
365	1.09 ± 0.02	0.83 ± 0.02	0.956‡	0.66 ± 0.01	—	—	—	1.55 ± 0.02	—	—	−9.5 ± 1	−7 ± 2
510	0.98 ± 0.02	0.95 ± 0.01	0.95 ± 0.01	0.69 ± 0.02	0.85 ± 0.02	—	0.92 ± 0.02	—	1.06 ± 0.03	0.8 ± 0.02	−2.5 ± 2	0 ± 1

Abbreviations: Gly, glycine; BSA, bovine serum albumin; Carb, carbonic anhydrase; Chym, α-chymotrypsinogen A; Myo, myoglobin; Aldo, aldolase; Cyt *c*, cytochrome *c*; Mk, myokinase (see ref. 3 for M_r and pI values). To determine R_i, a substantial mass (60–100 mg) of skinned muscle fibres (~1 g of skinned fibres could be obtained from a single barnacle) was initially equilibrated in the appropriate solution containing the substance *i* and then transferred after careful blotting to an 'unloading' solution of specially chosen composition to allow accurate determinations of the amount of *i* carried by the known mass of skinned muscle fibres into the unloading solutions. For protein determinations the unloading solutions had the composition listed in the Fig. 1 legend and for $EGTA^{2-}$ and K^+ determinations the unloading solutions contained only 200 mM KCl and 100 mM Tris-HCl (*p*H 7.10), respectively. The concentration of proteins in the loading solutions was 10 mg ml^{-1}. This concentration of proteins did not have any significant effect on R_K, R_{EGTA} or ΔV measurements. Similar results were obtained in control experiments in which the concentration of BSA was varied between 1 and 50 mg ml^{-1}. The density of the fibres was also determined in various solutions by dividing the mass of the fibres by the volume they displaced in these solutions. The initial mass of the muscle preparations mechanically skinned under paraffin oil increased by 40–50% following the detergent treatment and equilibration in solutions containing 100 mM K^+. Their mass increased further after equilibration in solutions with a higher ionic concentration[9]. The following analytical methods were used for the various R_i determinations: R_{EGTA}, titration method of Moisescu and Pusch[10]; R_{Gly}, spectrophotometric method using glycine reaction with ninhydrin; R_{protein}, trichloracetic acid precipitation followed by Lowry's method for protein determinations: $R_{Cyt\,c}$, both Lowry's method and direct spectrophotometric measurements; R_K, atomic absorption spectrophotometry.

* Composition of solutions as in Fig. 2 legend.

† $RT/F = 25.5$ mV.

‡ Interpolated values.

with Na^+, indicating the absence of a selective barrier, such as a membrane, between the two regions. An increase in $[Ca^{2+}]$ from $\sim 10^{-8}$ M to 10^{-5} M also did not result in any significant change in the potential differences recorded. However, an increase in $[Mg^{2+}]$ from ~0.2 to 5 mM resulted in a significant decrease in the potential difference between the two regions (~5 mV for solutions similar to those in Fig. 1). Relatively large potential differences could also be recorded in mechanically skinned amphibian muscle fibres (Fig. 1*C*) and in freshly skinned barnacle muscle fibres immersed in oil (Fig. 1*D*).

To characterize further the potential difference (ΔV) between the myofibrillar space and extra-myofibrillar space, we investigated the effect on ΔV of varying the concentration of KCl in the bathing solution. As can be seen from Fig. 2, ΔV decreases with an increase in the ionic concentration of the solutions in a fashion similar to that expected for a Donnan potential[2].

An important inference from the existence of a Donnan equilibrium between different regions of the muscle fibre is the non-uniform distribution of ions in this cell, and we have used the same skinned muscle preparation to determine the distribution of ions between the myofibrillar and extra-myofibrillar space. The extra-myofibrillar region of the skinned muscle fibre can be considered as an extension of the space surrounding the preparation which is in electrochemical (Donnan) equilibrium with the space between myofilaments. If there was an important difference in the ionic composition between the myofibrillar and extra-myofibrillar space due to a Donnan-type equilibrium produced by negatively charged myofilaments, the ratio, R_i of the average concentration of an ion *i* in the preparation to the concentration of the same ion in the bathing solution, would be expected to be significantly higher for cations than for anions (see Table 1 legend for methods used to determine R_i). This is confirmed by the results shown in Fig. 3, where the ratios for K^+ and $EGTA^{2-}$ have been obtained simultaneously in the same solutions and the same batch of skinned fibres from a single barnacle. The average values for R_K and R_{EGTA} obtained over all experiments are given in Table 1.

Figure 3 and Table 1 also show the R_i values for glycine, a small uncharged molecule, which would be expected to penetrate freely the space between myofilaments. The value of R_{Gly} can be used as an index of the fraction of total fibre volume available for diffusion (S_d), and these results indicate that most of the space within our preparations is free for diffusion in all the ionic conditions used. The fraction of total fibre volume which is extra-myofibrillar (S_e) can be estimated directly from the value of R_i of a relatively large, negatively charged protein which cannot enter the myofibrillar space and we have used bovine serum albumin (BSA, M_r 68,000, *p*I 5.89)[3] as a marker for this purpose. The values of R_{BSA} are shown in Table 1. If R_{BSA} is equated to S_e, it can be seen that the extra-myofibrillar space would represent more than 50% of the total volume of these skinned muscle preparations. Indeed, electron micrographs also indicate that large areas of extra-myofibrillar space exist in extensive parts of these Triton-treated skinned muscle fibres.

If the ions are in electrochemical equilibrium in the preparation, one can predict an average value for the electrical potential difference ($\Delta V^i_{\mathrm{pred}}$) between the myofibrillar and extra-myofibrillar space from the Nernst equation for each ion *i*. Using the experimentally determined values for R_i and assuming, for simplicity, that $R_{Gly} = S_d$, $R_{BSA} = S_e$ and the activity coefficients of the respective ions are similar in the myofibrillar

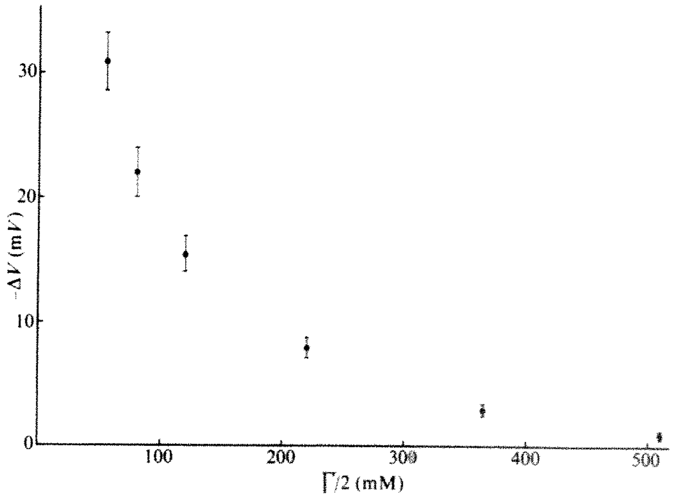


Fig. 2 Effect of ionic strength (Γ/2) on the average potential differences measured in mechanically skinned muscle fibres of the barnacle. The points represent the mean values (±s.e.m.) obtained from fibres of six different barnacles. Each fibre was impaled at least 30 times in each solution and care was taken to distribute the impalements over many regions of the fibre. The ionic strength was varied by adding differing amounts of KCl to a basic relaxing solution containing (mM): TES, 40; K_2EGTA, 10; Na_2ATP, 3; $MgCl_2$, 2. The final *p*H was adjusted with KOH to 7.10 in all solutions.

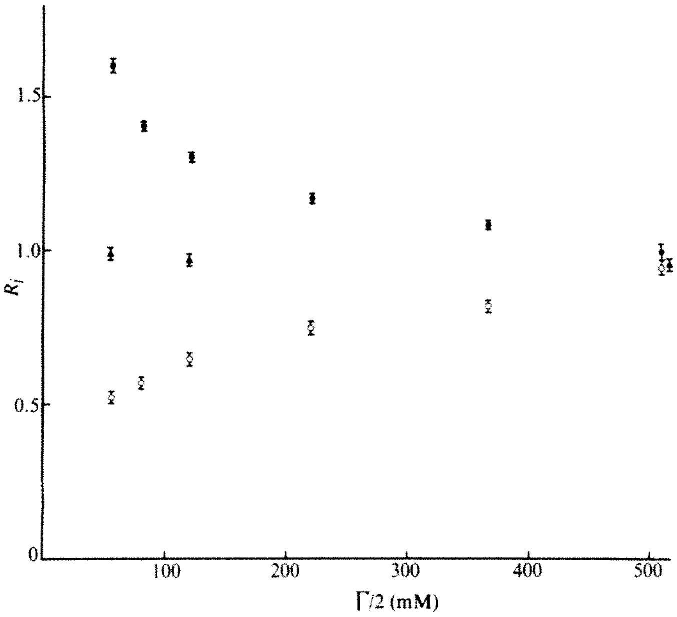


Fig. 3 Effect of ionic strength (Γ/2) on average values of R_{EGTA}(○), R_K(●) and R_{Gly}(▲), measured simultaneously in the same batch of skinned barnacle muscle fibres (200–300) from one animal. The vertical bars indicate the error of the methods (see Table 1). Note that R_{EGTA} is significantly smaller than R_K in all solutions but that of highest ionic strength and that the discrepancy between R_K and R_{EGTA} increases with decreasing ionic concentration in the solutions, as expected if a Donnan equilibrium was mainly responsible for the distribution of ions in these muscle preparations.

space and in the bulk solution, the predicted ranges for ΔV (ΔV^{K}_{pred} and ΔV^{EGTA}_{pred}) shown in Table 1 were obtained from the following equation

$$\Delta V^{i}_{pred} = (RT/n_iF)\ln(S_d - S_e)/(R_i - S_e)$$

In addition to BSA, the distribution of several other proteins of differing molecular weight and *p*I were investigated and their R_i values are listed in Table 1. The R_i values for the proteins of $M_r < 20{,}000$ approach the value for R_{Gly} as the ionic strength of the solutions increases, suggesting that there was no specific binding other than an electrostatic type of interaction between filaments and cytochrome *c*, or myoglobin, and that there was no steric restriction for the diffusion of these proteins between myofilaments. Cytochrome *c* and α-chymotrypsinogen A are both highly positively charged at *p*H 7.00 and show much larger values for R_i than does K^+ in conditions where a significant potential difference exists. Myoglobin is only slightly charged in these solutions and accordingly has an R_i value smaller than R_K.

The other proteins, carbonic anhydrase, aldolase and BSA, show a significantly smaller R_i value than R_{Gly} in the presence of 500 mM K^+ compared with $R_{EGTA} \simeq R_K \simeq R_{Gly}$, suggesting that these proteins are sterically restricted to certain degrees from entering the myofibrillar space. Of these proteins only aldolase is highly positively charged and if free to diffuse between myofilaments it would have been expected to have R_i values close to $R_{cyt\,c}$, which is not the case. However, the R_i value for aldolase is significantly higher than R_{BSA} for lower ionic strength solutions and this could be due to an increased concentration of aldolase in the outer layers of myofibrils due to the negative potential differences. These results on protein distribution: (1) strongly support the existence of a significant Donnan equilibrium within muscle fibres, (2) indicate that steric restrictions can become significant for the diffusion of proteins with $M_r >$ 20,000–30,000 and (3) suggest that the location of sarcoplasmic biochemical reactions would be determined by the charge and the size of the molecules involved. An interesting corollary from these experiments is that the prevailing electrical potential configuration within skeletal muscle fibres would promote a rapid activation of contraction by selectively facilitating the diffusion of Ca^{2+} released from the sarcoplasmic reticulum towards the area of the filament lattice with the highest negative charge density (region of filament overlap), where force is generated.

Previous attempts to record electrical potential differences between bundles of glycerinated muscle fibres[4–6] and the bathing solutions, and between mechanically skinned muscle fibres[7,8] and the bathing solutions have been reported. However, doubts have been raised about the state of membranes in the multicellular glycerinated preparations[6] and more recent observations on mechanically skinned muscle fibres indicated either a complete absence or drastically reduced potential differences[7,8] between the skinned muscle preparations and the bathing medium. The present results, obtained using several independent measurements, show that in the absence of any cellular membranes a significant electrical potential difference exists between regions within these muscle fibres and that electrically charged substances are not uniformly distributed within muscle fibres at physiological levels of monovalent cations.

We thank Mrs R. Cafarella for technical assistance and the Australian Research Grants Committee and National Heart Foundation of Australia for financial support.

Received 21 July; accepted 2 December 1980.

1. Ashley, C. C. & Moisescu, D. G. *J. Physiol., Lond.* **270,** 627–652 (1977).
2. Overbeek, J. M. G. *Prog. Biophys. molec. Biol.* **6,** 57–84 (1956).
3. Righetti, P. G. & Caravagio, T. *J. Chromat.* **127,** 1–30 (1976).
4. Collins, E. W. & Edwards, C. *Am. J. Physiol.* **221,** 1130–1133 (1971).
5. Pemrick, S. M. & Edwards, C. *J. gen. Physiol.* **64,** 551–567 (1974).
6. Weiss, R. M., Lazzara, R. & Hoffman, B. F. *Nature* **215,** 1305–1307 (1967).
7. Gordon, A. M., Godt, R. E., Donaldson, S. K. B. & Harris, C. E. *J. gen. Physiol.* **62,** 550–574 (1973).
8. Fabiato, A. & Fabiato, F. *Circulation Res.* **31,** 293–307 (1972).
9. Godt, R. E. & Maughan, D. W. *Biophys. J.* **19,** 103–116 (1977).
10. Moisescu, D. G. & Pusch, H. *Pflügers Arch. ges. Physiol.* **355,** R122 (1975).

Rhythmic contractile activity of the *in vivo* rabbit aorta

Allen Mangel, Mohammad Fahim
& Cornelis van Breemen

Department of Pharmacology, University of Miami School of
Medicine, PO Box 016189, Miami, Florida 33101

The cellular element of the aorta is largely smooth muscle; yet this organ has long been regarded as a passive elastic tube[1–4], without regard for the possible function of its smooth muscle cells. The isolated rabbit aorta became popular as a convenient smooth muscle preparation[5], and much has been learned about its cellular physiology, but unfortunately the functional role of this tissue was lost on its excision. Like many mammalian arterial preparations, the isolated aorta either fails to show spontaneous electrical or mechanical activity, or shows activity of a much slower frequency than the rhythmic activity of the heart[6,7], thus obscuring any possible relationship with the pulsatile activity of the heart. However, contractions are elicited from the isolated aorta on stimulation with various neural and hormonal agents[5–10]. In effect, the aortic smooth muscle has been classified as multiunit, which suggests that physiological activation may be neurogenic[11]. Thus, removal of the muscle from the body may result in disruption of the neural connections which function in normal activation of contraction. We show here that when recordings of aortic tension are made *in vivo*, rhythmic contractions are observed. Evidence is presented which indicates that the contractions are neurogenic in origin and exhibit a precise phasing pattern with the pulse wave.

A longitudinal cut was made along the abdominal midline of rabbits anaesthetized with an intravenous injection of Nembutal (35 mg per kg) through the ear vein. The abdominal aorta was exposed, and the blood flow from a 5-cm section was bypassed using a polyethylene tube (Fig. 1). In the centre of the bypass tube, a catheter was mounted and connected to a blood pressure transducer (Statham P23Db). An isometric tension transducer (Grass FTO3C) was attached to the bypassed muscle segment such that simultaneous pressure and tension recordings could be made. These recordings were made at loci equidistant from the heart.

Illustrated in Fig. 2 are simultaneous recordings of pulse pressure, measured in the bypass tube, and muscle tension, recorded from the bypassed muscle segment. The pressure waves and muscle contractions exhibited a 1 : 1 correspondence with the phasing between the two events, displaying a definite relationship. During the pressure wave diastole, muscle tension was maximum. Generally, within 20 ms after the initiation of an increase in pressure, the muscle began to relax and tension was usually at a minimum by a time corresponding to the dichrotic notch. As pressure began to decline again, muscle tension once again increased. This increase in tension was characterized by two components, a fast and a slow phase. The relative duration of the two phases varied somewhat from preparation to preparation. Such a cyclic pattern of simultaneous pressure and tension changes was observed in 10 of 10 animals tested.

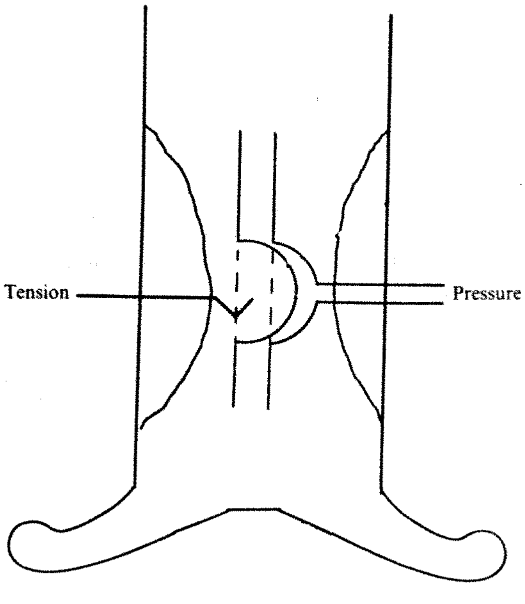


Fig. 1 Schematized recording technique for simultaneous measurement of pressure and tension in rabbit aorta. Dotted lines indicate bypassed muscle segment.

The possibility that the tension recordings were produced artefactually by body movements secondary to the pulse wave was eliminated by the lack of parallelism between the two recordings. Local application to the bypassed aortic muscle segment of the calcium flux inhibitor lanthanum (3 mM), local anaesthetics xylocaine (1%) and tetrodotoxin (10^{-7}–10^{-6} M), or the α-adrenergic blocker phentolamine (10^{-4} M) inhibited or abolished the contractions without affecting the pulse pressure. The contractions were enhanced by local application of atropine (10^{-4} M) or noradrenaline (10^{-6}–10^{-5} M), again without any effect on pulse pressure. If, on the other hand, noradrenaline (10^{-6} M) was administered parenterally through the ear vein, the pulse pressure increased while the contractions in the bypassed segment either decreased in amplitude or remained unchanged.

The abolition of the contractions by tetrodotoxin indicates that their propagation is neurally mediated. This premise was further investigated by measuring the direction and conduction velocity of the contractile wave. Figure 3 illustrates the effects of interrupting the anatomical continuity of the aorta on either side of the segment from which the contraction recordings were made. If the aorta was severed cranially (upstream), the contractions ceased completely (Fig. 3*b*), whereas after a distal cut contractions continued (Fig. 3*a*). The signal for contraction is therefore conducted in the same direction as the pulse wave. To measure the velocity of conduction, we bypassed two segments of the aorta and recorded contractions at two locations, the distance between which was accurately measured. The time interval between the onset of the contractions yielded a conduction velocity of 1.9 m s^{-1} directed caudally, a similar conduction velocity to that found for the pressure wave[12]. This velocity falls well within the range of neural activity[13,14] but is much faster than the propagation velocity observed for myogenic phenomena in other types of smooth muscle[15,16].

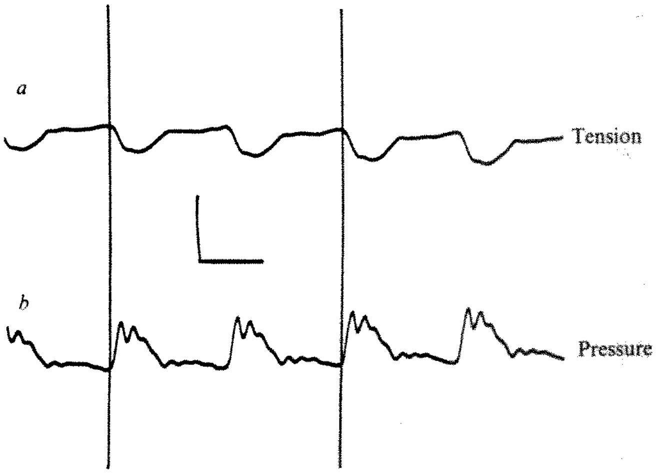


Fig. 2 Simultaneous recording of tension (*a*) and pressure (*b*) from the *in vivo* rabbit aorta. Time mark lines are drawn to coincide with the onset of the pressure wave. Scale bar: 0.6 g, 30 mm, 120 ms; basal blood pressure is 76.0 mm. An upward deflection reflects an increase in tension and pressure.

From the above results, we conclude that neural impulses activate waves of smooth muscle contraction, travelling with the

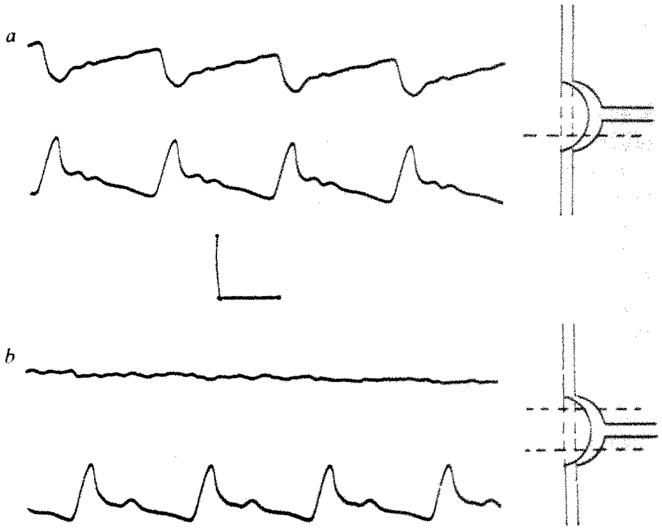


Fig. 3 Simultaneous recording of tension (upper) and pressure (lower) from the *in vivo* rabbit aorta. In (*a*), the distal side of the bypassed segment was severed from the intact aorta; in (*b*), the proximal side was then cut. Scale bar: 0.3 g, 40 mm, 150 ms; basal blood pressure in both (*a*) and (*b*) is 85 mm.

same velocity as the pulse pressure waves. The resulting synchrony between the pulse and contraction waves ensures that the aorta relaxes as pulse pressure increases and contracts during pulse pressure diastole.

The physiological function of the aortic oscillations remains a topic for conjecture. A previous study[17] has suggested that the aortic arch may undergo a contraction–relaxation cycle with a 1:1 correspondence with the heartbeat, possibly aiding in the diastolic propulsion of blood. Another role may be to reduce mechanical stress on the aortic wall. Prokop *et al.*[18] showed that the rate of change of blood pressure, but not its absolute magnitude, is the primary cause of aortic rupture. Thus, the observed relaxation during the increase in pulse pressure may locally reduce the rate of change of blood pressure and protect the aortic wall from mechanical damage.

This research was supported by American Heart Association grant 781124, NIH grants HL 07188 and HL 23932, and American Heart Fellowship RF 1. We thank Barbara Bradie for typing the manuscript and Gina del Valle for technical assistance.

Received 8 September; accepted 31 December 1980.

1. Bazett, H. C. *Am. J. Physiol.* **70,** 550–585 (1924).
2. Maltby, A. B. & Wigers, C. J. *Am. J. Physiol.* **100,** 604–613 (1932).
3. Hamilton, W. F. & Dow, P. *Am. J. Physiol.* **125,** 48–59 (1939).
4. Alexander, R. S. *Fedn Proc.* **11,** 738–749 (1952).
5. Furchgott, R. F. & Bhadrakom, S. *J. Pharmac. exp. Theor.* **108,** 129–143 (1953).
6. Biamino, G. & Kruckenberg, P. *Am. J. Physiol.* **217,** 376–352 (1969).
7. Altura, B. M. & Altura, B. T. *Microvascular Res.* **7,** 145–155 (1974).
8. Furchgott, R. F., Ehrreich, S. J. & Greenblatt, E. *J. gen. Physiol.* **44,** 499–519 (1961).
9. van Breemen, C. *et al. Phil. Trans. R. Soc.* **B265,** 57–71 (1973).
10. Deth, R. & van Breemen, C. *J. Membrane Biol.* **40,** 363–380 (1977).
11. Fischer, E. *Physiol. Rev.* **24,** 467–490 (1944).
12. Dow, P. & Hamilton, W. F. *Am. J. Physiol.* **125,** 60–65 (1939).
13. Mountcastle, V. B. (ed.) *Medical Physiology* Vol. 1, 13th edn (Mosby, St Louis, 1974).
14. Ganong, W. F. *Review of Medical Physiology* 2nd edn (Lange Medical, Los Altos, 1969).
15. Tomita, T. *Prog. Biophys. molec. Biol.* **30,** 185–202 (1975).
16. Connor, J. A., Mangel, A. M. & Nelson, B. *Am. J. Physiol.* **237,** C237–C246 (1979).
17. Boucek, R. J., Takashita, R. & Fojaco, R. *Am. J. Anat.* **114,** 273–282 (1964).
18. Prokop, E. K., Palmer, R. F. & Wheat, M. W. *Circulation Res.* **27,** 121–127 (1970).

No need for a new membrane model

Fritz Jähnig

Max-Planck-Institut für Biologie, Correnssstrasse 38,
74 Tübingen, FRG

Schindler *et al.*[1] recently reported lateral diffusion measurements in reconstituted membranes of phospholipid (PL), lipopolysaccharide (LPS) and *Escherichia coli* matrix protein (P), using the technique of fluorescence recovery after photobleaching (FRAP). Evaluation of their data led Schindler *et al.* to conclude that the fluid mosaic model is an inadequate description of the membrane and to propose a new membrane model. Their conclusion was based on identifying the fluid mosaic model with a particular binding behaviour of LPS to matrix protein. I present here a more general model for the association of membrane components, and demonstrate the use of lateral diffusion data in elucidating membrane structure. The data of Schindler *et al.* are shown to be reasonably interpretable on the basis of an association of LPS and matrix protein, which obviates the necessity for postulating a new membrane model.

I begin with a discussion of the lateral diffusion in a mixed membrane observed by the FRAP technique. Schindler *et al.*[1] studied the lateral diffusion constant D_S of one component such as LPS or PL, termed substrate, at different protein concentrations. The substrate molecules may be subdivided into those which are bound to protein and do not exchange in the time scale of a FRAP measurement, and the rest. D_S is given by the contributions from the non-exchanging molecules of molar fraction n_{nonex} and the remainder of the molar fraction n_{ex}: $D_S = D_P\, n_{\text{nonex}} + D_{\text{ex}}\, n_{\text{ex}}$. The diffusion constant of the non-exchanging molecules equals the protein diffusion constant D_P, which is usually much smaller than the diffusion constant of LPS or PL. If D_P vanishes, the fraction n_{nonex} of the substrate molecules does not diffuse, and the fluorescence recovery r is $<100\%$ and given by $r = n_{\text{ex}}$. In the measurements of Schindler *et al.*[1] the PL recovery was 100% and the LPS recovery >75%, implying that some LPS molecules were fixed to matrix protein and that the matrix protein diffusion was negligibly small. D_P was indeed found to be $<10^{-12}\ \text{cm}^2\,\text{s}^{-1}$. In my treatment I will assume for simplicity that non-exchanging molecules do not exist: $n_{\text{nonex}} = 0$ and $n_{\text{ex}} = 1$, so that $D_S = D_{\text{ex}}$.

The molecules denoted as exchanging may diffuse completely freely or may interact with proteins. In the latter case, they may bind to proteins and thus become immobilized for some time, but then exchange with free molecules. Assuming the exchange to be rapid on the time scale of a FRAP experiment ($t > 1$ s), the diffusion constant D_{ex}, which is the temporal average over the free and bound states of a molecule, equals an ensemble average over the exchanging molecules at any time: $D_{\text{ex}} = D_f n_f + D_P n_b$. Here D_f is the diffusion constant for free diffusion, and the diffusion constant of the bound molecules is again equal to the protein diffusion constant D_P; n_f and n_b are the molar fractions of free and bound molecules with $n_f + n_b = 1$. I assume $D_P = 0$. Furthermore, D_f is regarded as independent of protein concentration, which is valid only if proteins do not affect the viscosity of the free-substrate phase. I introduce a normalized diffusion constant $\tilde{D}_S(\text{P}) = D_S(\text{P})/D_S(0)$, which is then simply given by the concentration of free substrate:

$$\tilde{D}_S(\text{P}) = n_f(\text{P}) \qquad (1)$$

The function $n_f(\text{P})$, a static quantity reflecting association of substrate and protein, can be derived from the law of mass action. We assume a simple binding equilibrium $\text{S} + \text{P}' \rightleftarrows \text{SP}'$, where P' denotes protein-binding sites. The dissociation constant K_d is thus

$$K_d = \frac{\text{S}_f \text{P}'_f}{\text{S}_b} \qquad (2)$$

The concentration of free binding sites is given by $\text{P}'_f = \text{P}'_{\text{tot}} - \text{S}_b$, where P'_{tot} is the total concentration of binding sites. The latter equals $b\text{P}$, with b representing the number of binding sites per protein molecule. Equation (2) can be transformed into

$$\text{S}_b = b\text{P}\, \frac{\text{S}_f}{K_d + \text{S}_f} \qquad (3)$$

Insertion of $\text{S}_b = \text{S} - \text{S}_f$, where S is the total substrate concentration, produces a quadratic equation for S_f which can be solved to yield the following expression for $n_f = \text{S}_f/\text{S}$:

$$n_f = \tfrac{1}{2}(1 - b\text{P}/\text{S} - K_d/\text{S})\left(1 \pm \left[1 + \frac{4K_d/\text{S}}{(1 - b\text{P}/\text{S} - K_d/\text{S})^2}\right]^{1/2}\right) \qquad (4)$$

The sign has to be chosen so that $n_f \geq 0$. Note that n_f depends on the protein concentration via the molar ratio P/S, which will therefore be used as the measure of protein concentration. Equations (1) and (4) represent the variation of the lateral diffusion of LPS or PL with protein concentration.

To resolve further this result we discuss two limiting cases: first, the case of small protein concentration $b\text{P}/\text{S} \ll 1$. Here equation (4) becomes

$$n_f = 1 - \frac{b\text{P}}{\text{S}}\, \frac{\text{S}}{K_d + \text{S}} \qquad (5)$$

which is a functional relationship well known from enzyme kinetics. It describes the saturation of a few protein-binding sites with substrate where the number of bound molecules is proportional to the protein concentration. This implies that the substrate diffusion constant $\tilde{D}_S$ decreases linearly with protein

0028-0836/81/080694—03$01.00

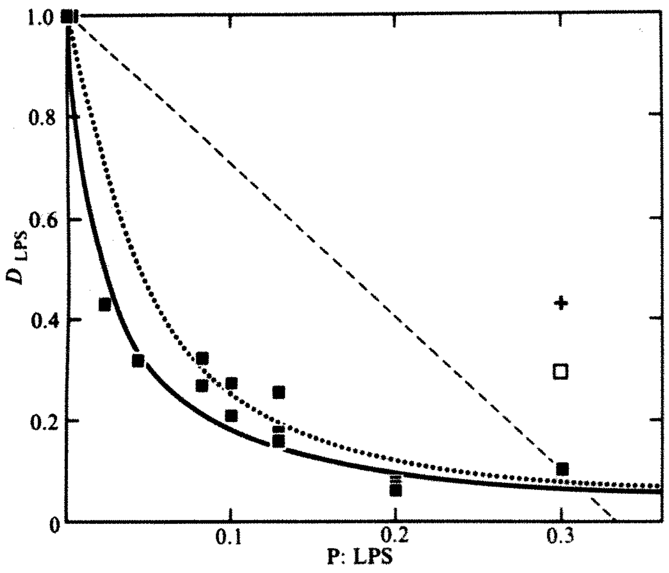


Fig. 1 Normalized lateral diffusion coefficient $\tilde{D}_{\rm LPS}$ of LPS in membranes of PL, LPS and matrix protein as a function of the molar ratio P/LPS. Experimental data of Schindler *et al.*[1] for molar ratios PL/LPS of 6 (■) and 72 (□); for the conversion of weight % protein and weight ratio PL/LPS (given by Schindler *et al.*[1]) into molar ratios, the molecular weights $\mathrm{MW_P} = 36{,}000$, $\mathrm{MW_{LPS}} = 3{,}600$ and $\mathrm{MW_{PL}} = 600$ were used. Theoretical curves are shown for PL/LPS = 6 in the limit of high P, equation (6), with $\kappa b = 325$ (—); in the general case, equation (4), with $\kappa b = 325$ and $b = 21$ (···); and in the limit of low P, equation (5), with $K_{\rm d} = 0$ and $b = 3$ (---); and a theoretical point for PL/LPS = 72 in the limit of high P with $\kappa b = 325$(+).

concentration at constant substrate concentration. The second limiting case is that of high protein concentration $b\mathrm{P/S} \gg 1$, where equation (4) yields

$$n_{\rm f} = \frac{K_{\rm d}}{K_{\rm d} + b\mathrm{P}} \tag{6}$$

This describes the situation where an excess of protein binds the few substrate molecules, but only succeeds in binding all of them asymptotically at infinite protein concentration. In this case there is saturation of the few substrate molecules by protein, which is the opposite of the first case. Here the substrate diffusion constant $\tilde{D}_{\rm S}$ shows an asymptotic decrease to zero with increasing protein concentration.

Thus in this simple binding model there is a linear decrease in the substrate diffusion constant at low protein concentration, due to protein saturation, and a gradual decrease to zero at high protein concentration, due to substrate saturation. Experimental results for $\tilde{D}_{\rm S}$ can be evaluated to yield the number b of protein-binding sites and the dissociation constant $K_{\rm d}$.

The variation of $\tilde{D}_{\rm S}$ with protein and substrate concentration derived above can also be formulated in another model which includes a partition coefficient. Two possible advantages of such a model are: first, it may be more appropriate in the case of many protein binding sites, $b \gg 1$; second, it may lend itself to further generalization concerning interactions between substrate and membrane components other than protein. Consider the partitioning of a substrate molecule between protein and lipid consisting of PL and LPS: the probability $p_{\rm P}$ that a substrate molecule is bound to protein is proportional to the concentration of free protein-binding sites (P′), and the probability $p_{\rm L}$ of association with lipid is proportional to the concentration of free lipid sites (L), so that

$$\frac{p_{\rm P}}{p_{\rm L}} = \frac{\mathrm{P}'_{\rm tot} - \mathrm{P}'_b}{\mathrm{L}_{\rm tot} - \mathrm{L}_b} \cdot \kappa \tag{7}$$

where κ is the partition coefficient. The substrate molecules in the lipid phase have been denoted free, therefore $p_{\rm P}/p_{\rm L} = \mathrm{S}_b/\mathrm{S}_{\rm f}$ at equilibrium. On the right-hand side of equation (7) we can assume $\mathrm{L_{tot}} \gg \mathrm{L}_b$, and substitute $b\mathrm{P}$ for $\mathrm{P}'_{\rm tot}$ and S_b for P'_b. Solving for S_b we obtain a relation which is identical to equation (3) and has the same solution equation (4), with the substitution

$$K_{\rm d} = \mathrm{L_{tot}}/\kappa \tag{8}$$

Thus the partition coefficient model leads back to the binding model. However, due to the appearance of the lipid-binding sites, $\mathrm{L_{tot}}$, the partition coefficient model can also account explictly for the interaction of substrate, for example LPS, with PL and LPS. The simplest approximation is

$$\mathrm{L_{tot}} = \mathrm{PL} + \mathrm{LPS} \tag{9}$$

The final result for $\tilde{D}_{\rm S}$ is obtained in the form of equation (4) for $n_{\rm f}$ with the substitution of $K_{\rm d}$ according to equations (8) and (9). In the limiting case of high protein concentration as in equation (6), the following is obtained:

$$\tilde{D}_{\rm S}\left(\frac{\mathrm{P}}{\mathrm{LPS}}, \frac{\mathrm{PL}}{\mathrm{LPS}}\right) = 1\Big/\left(1 + \frac{\kappa b}{1 + \mathrm{PL/LPS}}\,\frac{\mathrm{P}}{\mathrm{LPS}}\right) \tag{10}$$

These theoretical results can now be used to evaluate the experimental data of Schindler *et al.*[1] for the lateral diffusion of LPS. They are presented as $\tilde{D}_{\rm LPS}$ (P/LPS) in Fig. 1. The data points for PL/LPS = 6 show a gradual decrease to zero with increasing protein concentration indicating the limiting case of high protein concentration. To test this suggestion we plotted $\tilde{D}^{-1}_{\rm LPS}$ against P/LPS (see Fig. 2) which should yield a straight line according to equation (10). Although the data points scatter appreciably, they can be regarded as conforming to a straight line; a least-squares fit gives $\kappa b = 325$. There is no indication of a cross-over of the experimental points from the limit of high protein concentration to the case of low protein concentration, equation (5); therefore the quantitites b and κ (or $K_{\rm d}$) cannot be determined separately. A lower limit for b may be estimated from the condition of high protein concentration, $b\mathrm{P/LPS} \gg 1$. Assuming that this condition conforms to the experimental point at P/LPS = 0.047 in Fig. 2, we obtain $b \gg 21$. The theoretical curve for $\tilde{D}_{\rm LPS}$ with $b = 21$, derived from the general

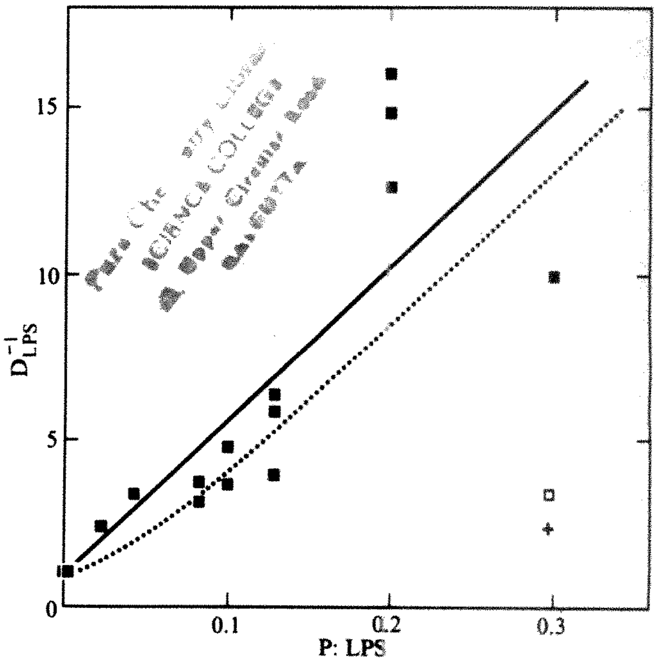


Fig. 2 Inverse of the normalized lateral diffusion coefficient $\tilde{D}_{\rm LPS}$. Representation of points and lines as in Fig. 1.

expression (4), is included in Figs 1 and 2. Comparison with the earlier fit based on equation (10), which may be regarded as the limiting case for large b, shows that $b = 21$ represents a lower limit for b. Such a high number of binding sites does not seem unlikely because such binding is only a loose association between LPS and protein, comparable with lipid–protein association, where even higher numbers of boundary lipids associated with protein are known.

In contrast to the evaluation presented here, Schindler *et al.*[1] restricted themselves to $b = 3$ and $K_d = 0$. With these values equation (6) yields $\tilde{D}_{LPS} = n_f = 0$ at high protein concentration, and at low protein concentration equation (5) yields a straight line for $\tilde{D}_{LPS}$ (P/LPS). The latter is included in Fig. 1, showing its serious disagreement with the experimental data, which led Schindler *et al.*[1] to propose a new membrane model. They adopted the number of three binding sites from studies of the *E. coli* outer membrane, where it was found that three LPS molecules are associated with one matrix protein[2]. However, the assumption $b = 3$ would imply that the matrix proteins are saturated with LPS at the relatively high protein concentration of the outer membrane, which seems rather unlikely. Within our evaluation we would predict from equation (10), with $\kappa b = 325$ and the outer membrane molar ratios $P/LPS = 1/3$ and $PL/LPS = 6$, that $n_f = 0.06$, or 2.82 LPS molecules are associated with 1 matrix protein. This number for the site occupation agrees well with the experimental result, while the number of binding sites is expected to be much higher, as discussed. Note, however, that the evaluation of Schindler *et al.* and the one presented here differ not only in the number of binding sites, but also in the binding constant, producing a different dependence of $\tilde{D}_{LPS}$ on protein concentration.

Because the evaluation of Schindler *et al.* is too restricted, their conclusions concerning old and new membrane models are invalid. However, I hope I have made it clear that lateral diffusion measurements such as those of Schindler *et al.* are useful in studying the association behaviour of membrane constituents.

I thank many of my colleagues for stimulating discussions.

Received 30 June; accepted 1 December 1980.

1. Schindler, M., Osborn, M. J. & Koppel, D. E. *Nature* **283**, 346–350 (1980).
2. Schindler, H. & Rosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3751–3755 (1978).

Reply from D. E. Koppel*, M. J. Osborn† & M. Schindler‡

Departments of Biochemistry* and Microbiology†,
University of Connecticut Health Center, Farmington,
Connecticut 06032
‡ Department of Biochemistry, Michigan State University,
East Lansing, Michigan 48823

Some comments are in order to put Dr Jähnig's re-evaluation of our data in a proper perspective. We determined the lateral diffusion coefficients of lipopolysaccharide (LPS), phospholipid (PL) and *Escherichia coli* matrix porin protein in reconstituted multibilayer membranes[1], using the technique of fluorescence redistribution after photobleaching. We compared the results from a series of samples of constant relative LPS and PL content (1:1 by weight), but different protein content (0–60% by weight). It was found that as the protein concentration increased, the diffusion coefficient of LPS decreased by an order of magnitude, whereas that of PL decreased to a much smaller extent. The protein molecules at the higher concentrations were essentially 'immobile', with little or no redistribution observable over the distance scale (~3 μm) and maximum time scale (~3 × 10^3 s) of these experiments. This immobility would be a surprising result in the context of the fluid mosaic model, but is not unexpected in view of considerable evidence demonstrating the self-association of porin trimers into ordered aggregates *in vivo* as well as *in vitro*[2].

Jähnig has chosen to interpret these results solely in terms of specific LPS binding to the immobile protein, leaving both the dissociation constant and the stoichiometry of binding as floating parameters. This analysis can yield results consistent with the data, but only if one allows a total of at least 21 LPS binding sites per protein monomer (that is, ~63 sites per trimer).

In contrast to the above, we[1] did not assume that in the absence of specific binding the diffusion rates would be necessarily unaffected by the presence of an immobile protein matrix in the membrane, but set out to try to determine what other factor, if any, could be contributing to the observed diffusion characteristics. We evaluated the possible effect of LPS binding consistent with that which we considered to be the best available value of binding stoichiometry. Based on evidence of copurification, we estimate this as 9 LPS molecules per protein trimer[3,4]. Under this constraint, the binding model of diffusion could not account for the magnitude of the observed decrease in LPS diffusion. As a possible explanation of this difference we speculated that the translational diffusion of membrane components in these conditions might bear a closer resemblance to diffusion through a polymeric network than to diffusion in a simple viscous fluid. The differential effects observed for LPS and PL would be explained by the different sizes of the diffusing species. The part of the polymeric network in this scenario is played by the immobile protein matrix. Subsequent experiments[5] have suggested that the peripheral protein matrix of erythrocyte membranes may exert a similar control over the diffusion of integral membrane proteins projecting beyond the membrane bilayer.

Evidence is not yet available to allow a final definitive analysis. The value of the stoichiometry we have selected is, admittedly, a relatively soft number—it could be higher. At the same time, the value deduced by Jähnig seems unreasonably high for a specific binding interaction. The truth probably lies somewhere in between.

1. Schindler, M., Osborn, M. J. & Koppel, D. E. *Nature* **283**, 346–350 (1980).
2. Osborn, M. J. & Wu, H. C. P. *A. Rev. Microbiol.* **34**, 369–422 (1980).
3. Schindler, H. & Rosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3751–3755 (1978).
4. Rosenbusch, J. P., Steven, A., Alkan, M. & Regenass, M. in *Regular 2-D Arrays of Biomolecules: Structure Determination and Assembly* (ed. Baumeister, W.) (Springer, Berlin, 1979).
5. Sheetz, M. P., Schindler, M. & Koppel, D. E. *Nature* **285**, 510–512 (1980).

Molecular cloning of the K1 capsular polysaccharide genes of *E. coli*

Richard P. Silver*, Charles W. Finn*, Willie F. Vann*, Wendy Aaronson*, Rachel Schneerson*, Peter J. Kretschmer†§ & Claude F. Garon‡

* Division of Bacterial Products, Bureau of Biologics, FDA,
Bethesda, Maryland 20205
† Molecular Hematology Laboratory, National Heart, Lung and Blood Institute, NIH, Bethesda, Maryland 20205
‡ Laboratory of Biology of Viruses, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland 20205

Epidemiological and immunological evidence indicates that the K1 capsular polysaccharide confers the property of virulence on *Escherichia coli*. *E. coli* K1 is associated with invasive diseases in humans and in laboratory and domesticated animals[1]. K1 isolates account for 80% of *E. coli* neonatal meningitis and comprise the majority of capsular types in neonatal septicaemia without meningitis and in childhood pyelonephritis[2,3]. Passive administration of K1 antibodies prevented bacteraemia and meningitis in infant rats fed *E. coli* K1[4]. Nonencapsulated

---

§ Present address: Bethesda Research Laboratories, Rockville, Maryland 20850.

derivatives of these invasive K1 strains did not cause bacteraemia in infant rats, although intestinal colonization was similar to that of the parent strains (M. Achtman and R.P.S., unpublished results). Several reports propose that the *E. coli* K1 capsular polysaccharide exerts an anti-phagocytic effect similar to that observed with other pathogenic encapsulated bacteria[5,6]. One approach to studying whether the K1 antigen is sufficient to confer virulence or if other *E. coli* structures are necessary is to isolate the K1 genes for genetic and biochemical analysis. Recombinant DNA methodology provides a powerful tool for such an approach. Here, we report the molecular cloning of the *E. coli* K1 antigen genes. The cloned K1 genes synthesize a capsule in *E. coli* K12 indistinguishable chemically and immunologically from that of wild-type K1 strains.

The K1 capsular polysaccharide is a linear homopolymer of α2–8-linked *N*-acetylneuraminic acid (NANA)[7,8]. A genetic locus, associated with K1 biosynthesis, linked to the *serA* locus on the *E. coli* chromosome and termed *kpsA*, has been described elsewhere[9]. A membrane-associated sialyltransferase complex catalyses the incorporation of sialic acid from CMP-sialic acid into polymeric products[10]. The number and function of the genes involved and the molecular events controlling the synthesis and assembly of the K1 polysaccharide have not been elucidated. Therefore, we chose to use the cosmid cloning vector pHC79 in conjunction with *in vitro* λ packaging techniques to clone the K1 genes[11,12]. The plasmid pHC79, a 6.4-kilobase derivative of pBR322, confers resistance to ampicillin (Amp) and tetracycline (Tc)[11]. *E. coli* RS218 (ref. 13) (*O*18, K1, H7) chromosomal DNA[14] (75 μg ml^{-1}), partially digested with *Bam*HI, and *Bam*HI-cleaved pHC79 (225 μg ml^{-1}), were mixed and ligated in a total volume of 20 μl (ref. 12). *Bam*HI inactivates the Tc resistance gene of pHC79. Ligated DNA (8 μl) was packaged in 400 μl of packaging lysate[12]. *E. coli* K12 strain HB101 (ref. 15) (*hsd*R, *hsd*M, RecA, *leu pro*) was transduced by diluting the packaging mix to 2.0 ml with L-broth[16](LB) (containing 200 mM $MgSO_4$) and adding the mixture to 10 ml of cells (10^9 per ml in LB grown in 0.4% maltose). After 10 min adsorption at 30 °C the mixture was diluted five-fold with LB, incubated for 1 h at 37 °C, concentrated 10-fold and plated on L-agar[16] or antiserum agar[17] containing 50 μg ml^{-1} ampicillin. Resistance to Tc was determined by growth on L-agar containing 20 μg ml^{-1} Tc. Complementation of proline auxotrophy was detected by the ability of HB101-containing hybrid plasmids to grow in minimal medium supplemented with leucine but lacking proline.

Amp-resistant transductants (4×10^4 per μg ligated *E. coli* DNA) were obtained from the L-agar plates. Thirty per cent of the transductants were Tc sensitive and 1.4% were proline prototrophs, confirming the insertion and expression of *E. coli* chromosomal DNA into the *Bam*HI site of pHC79. Amp^R colonies (6×10^3) were pooled from L-agar plates and stored at −20 °C. These cells were later grown in LB and their plasmid DNA isolated. Contour length measurements[18] of purified plasmid DNA showed a homogeneous population of 43.8 kilobases (14.78 ± 0.57 μm). This is consistent with sizes previously reported for cosmid hybrids[11].

The K1 antigen in the transduced cells was detected by halo (precipitin) formation on antiserum agar plates containing equine meningococcal group B antiserum (the K1 capsular polysaccharide is structurally and antigenically identical to the meningococcal group B polysaccharide)[8]. There were fewer Amp^R transductants per μg ligated *E. coli* DNA on antiserum agar than on L-agar and no halo-producing colonies were observed among more than 1,500 transductants. In contrast, when Amp^R cells from the L-agar colony pool were plated on antiserum agar, many halo-producing colonies were isolated. Our inability to isolate transductants containing K1, except after plating on medium lacking antiserum, may reflect a delay in the expression of the newly acquired K1 genes. *E. coli* K12 was sensitive to the bactericidal effects of the antiserum; Table 1 shows a 10^5-fold decrease in plating efficiency of HB101 on antiserum agar. The bactericidal effect was eliminated by heating the serum at 56 °C for 30 min, suggesting that serum sensitivity was due to complement proteins. *E. coli* K12 expressing the K1 antigen was not serum sensitive (Table 1). This is probably due to shielding effects of the capsule on deeper cell structures capable of activating complement proteins.

Another factor contributing to the difficulty in isolating K1 transductants was the instability of pHC79/K1 hybrid plasmids. We observed a 10^4-fold decrease in the number of K1$^+$Amp^R organisms following overnight growth in the absence of antibiotic selection. The instability of the plasmid is consistent with previous observations. Henning *et al.*[19] failed to find the *OmpA* gene in the *E. coli* hybrid gene bank of Clark and Carbon[20]. Nishimira *et al.*[21] also failed to find complementation of other outer membrane protein mutants in the Clark and Carbon[20] collection including the *OmpC* and lipoprotein genes. Henning *et al.*[19] were unable to clone the *OmpA* gene on a multicopy vector; *OmpA* was cloned on λ and pSC101 vectors. These data suggest that the gene dosage of outer membrane proteins and other surface structures, generated by cloning on multicopy

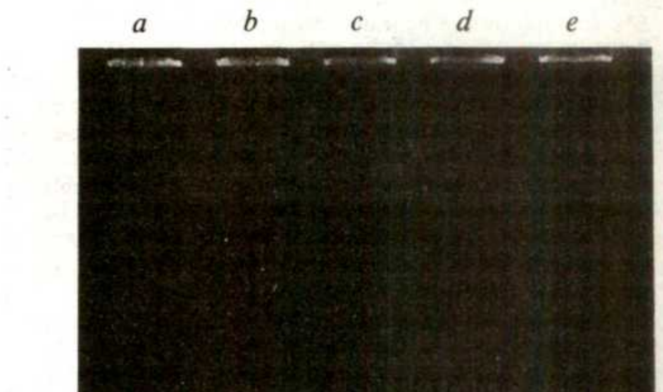


Fig. 1 Electrophoresis in 0.7% agarose gel of *Bam*HI-digested plasmids. Plasmid DNA was isolated as described elsewhere[13] and treated with endonuclease *Bam*HI in 100 mM Tris-HCl (*p*H 7.5), 10 mM $MgCl_2$ and 50 mM NaCl at 37 °C for 1 h. Samples were run in 0.7% agarose gels, Tris-borate buffer, *p*H 8.2, for 3 h at 120 V. DNA fragments were visualized by staining with ethidium bromide (0.4 μg ml^{-1}) and photographed in short-wavelength UV light. Molecular size standards of *Eco*RI-cleaved λ DNA were used to indicate fragment length. Lane *a*, pHC79; lane *b*, pSR23; lane *c*, pSR25; lane *d*, pSR26; lane *e*, pSR27. pSR25–27 were isolated following *Bam*HI digestion of pSR23, re-ligation[12] and transformation[22] of HB101. kb, Kilobases.

plasmid vectors, may be directly related to the lethality of the cell[19]. A recent observation that a strain in which the K1 genes have been transposed to a low copy number plasmid (F*lac*Mu) stably produced the K1 antigen is consistent with this view (R.P.S., unpublished observation).

The production of the K1 antigen could be maintained in broth cultures or on agar plates containing 500 or 100 μg ml^{-1} ampicillin, respectively. A 40-kilobase (13.7 ± 0.48 μm) plasmid, designated pSR23, was isolated from HB101Kl$^+$ cells. A *Bam*HI digest of pSR23 generated four fragments: the 6.4 vector and three *E. coli* fragments of 23, 6.6 and 5.0 kilobases (Fig. 1). A DNA–DNA heteroduplex between pHC79 and pSR23 confirmed the presence of a 35-kilobase insert in pHC79 (Fig. 2). HB101 was readily transformed[22] to ampicillin resistance by pSR23; all Amp^R transformants produced the K1 antigen. The efficiency of plating of K1 specific phage[23] was the same on HB101(pSR23) as on RS218. The 35-kilobase cloned *E. coli* fragment represents about 1 min of the *E. coli* genetic map[24]. It contains information to code for about 50 proteins and is larger than that required for K1 biosynthesis alone. Other *E. coli* genes on the hybrid plasmid have not been identified. Subclones containing pHC79 and each of the *Bam*HI fragments of pSR23 (Fig. 1, lanes *c–e*) did not produce K1 polysaccharide as determined by halo formation on antiserum agar and K1 phage sensitivity. Cells harbouring pSR27, which contains the largest *Bam*HI fragment (Fig. 1, lane *e*), synthesized NANA and had sialyltransferase activity (W.F.V. and R.P.S., unpublished results).

The structure of the K1 polysaccharide isolated from HB101(pSR23) was studied by several techniques[25]. We found that: (1) it was composed of 88% sialic acid; (2) the molecular size, determined by gel filtration, was similar ($k_d = 0.5$) to that of K1 polysaccharides isolated from other K1 strains; (3) the purified polysaccharide was negatively charged at neutral *p*H; (4) in double immunodiffusion a line of identity was observed between the polysaccharide from HB101(pSR23) and wild-type K1; and (5) the ^{13}C-NMR spectrum of the HB101(pSR23) polysaccharide was identical to that of wild-type K1.

Isolation of the K1 genes will permit a direct examination of whether the K1 capsular polysaccharide alone confers the property of invasiveness to *E. coli*. Our preliminary experiments indicate that *E. coli* K12 strains with the K1 capsular polysaccharide genes are not invasive for the infant rat. Thus, it is likely that biochemical traits in addition to the K1 capsule are required for the invasiveness of *E. coli*. The techniques of molecular biology and bacterial genetics should provide useful approaches to characterizing these pathogenic traits.

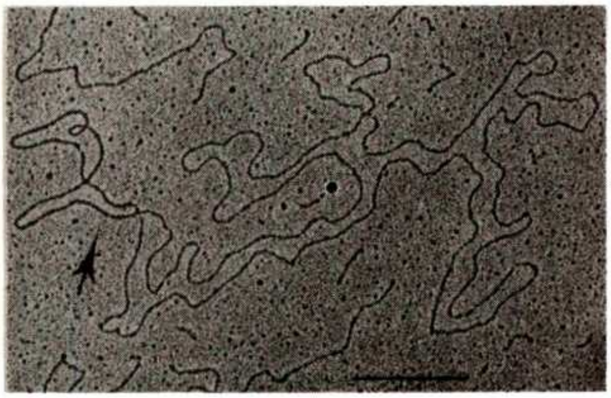


Fig. 2 Electron micrograph of DNA–DNA heteroduplex between pHC79 and pSR23. pHC79 was digested with endonuclease *Pst*I at 37 °C in 20 mM Tris-HCl (*p*H 7.5), 10 mM $MgCl_2$, 50 mM $(NH_4)_2SO_4$ generating linear molecules. pSR23 was incubated with pancreatic DNase I (0.02 units ml^{-1}) for 25 min at 25 °C, conditions producing approximately one nick per molecule. Heteroduplex analysis was performed as described previously[18]. Arrow indicates double-stranded 6.4-kilobase pHC79. Scale bar, 1 μm.

Table 1 Growth of *Escherichia coli* on antiserum agar

Strain	Plating efficiency
RS218	1.0
HB101	10^{-5}
HB101 (serum heated 56 °C, 30 min)	1.0
CSH75*	10^{-4}–10^{-5}
CSH75K1$^+$†	0.5

Cells were grown overnight in LB at 37 °C, diluted and plated on antiserum agar and L-agar. Efficiency of plating on antiserum agar was expressed relative to growth on L-agar.

* *E. coli* K12 strain (*ara, leu, lacY, purE gal trp his, argG malA, strA, xyl, mtl, ilv, met, thi, pro*).

† Chromosomal recombinant obtained from a mating with an *E. coli* K1 Hfr donor (R.P.S. and J. Foulds, unpublished).

We thank John B. Robbins and M. Carolyn Hardegree for reviewing the manuscript, William Egan for NMR studies, John Collins for providing pHC79, B. Rowe for providing K1 specific phage, and Helen C. Coon for assistance in preparation of the packaging extract.

Received 29 August; accepted 3 December 1980.

1. Schiffer, M. S. *et al. Pediat. Res.* **10,** 82–87 (1976).
2. Robbins, J. B. *et al. New Engl. J. Med.* **290,** 1216–1220 (1974).
3. Kaijser, B., Hanson, L. A., Jodal, V., Linden-Janson, G. & Robbins, J. B. *Lancet* **i,** 663–664 (1977).
4. Glode, M., Sutton, A., Moxon, R. & Robbins, J. B. *Infect. Immun.* **16,** 75–80 (1977).
5. Bortolussi, R., Ferrieri, P., Bjorksten, B. & Quie, P. G. *Infect. Immun.* **25,** 293–298 (1979).
6. Weinstein, R. & Young, L. S. *J. clin. Microbiol.* **8,** 248–255 (1978).
7. Barry, G. T. *J. exp. Med.* **107,** 507–521 (1959).
8. Kasper, D. L., Winkelhake, J. L., Zollinger, W. D., Brandt, B. L. & Artenstein, M. S. *J. Immun.* **110,** 262–268 (1973).
9. Orskov, I., Sharma, V. & Orskov, F. *Acta path. microbiol. scand.* B**84,** 125–131 (1976).
10. Troy, F. A. *A. Rev. Microbiol.* **33,** 519–560 (1979).
11. Collins, J. *Meth. Enzym.* **68,** 309–326 (1979).
12. Kretschmer, P. J. *et al. J. biol. Chem.* **255,** 3204–3211 (1980).
13. Silver, R. P., Aaronson, W., Sutton, A. & Schneerson, R. *Infect. Immun.* **29,** 200–206 (1980).
14. Smith, M. G. *Meth. Enzym.* **12,** 545–550 (1967).
15. Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41,** 459–472 (1969).
16. Miller, J. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
17. Sarff, L. D. *et al. Lancet* **i,** 1099–1104 (1975).
18. Davis, R. W., Simon, M. & Davidson, N. *Meth. Enzym.* **21,** 413–428 (1971).
19. Henning, V., Royer, H-D., Teather, R. M., Hindennach, I. & Hollenberg, C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4360–4364 (1979).
20. Clarke, L. & Carbon, J. *Cell* **9,** 91–99 (1976).
21. Nishimura, Y. *et al. Plasmid* **1,** 67–74 (1977).
22. Cohen, S. N., Chang, A. C. Y. & Hsu, L. *Proc. natn. Acad. Sci. U.S.A.* **69,** 2110–2114 (1972).
23. Gross, R. J., Cheasty, T. & Rowe, E. *J. clin. Microbiol.* **6,** 548–550 (1977).
24. Bachmann, B. J. & Low, K. B. *Microbiol. Rev.* **44,** 1–56 (1980).
25. Ørskov, F. *et al. J. exp. Med.* **149,** 669–685 (1979).

Erratum

In the letter 'Relative departures in July temperatures in northern Canada for the past 6,000 yr' by J. T. Andrews *et al.*, *Nature* **289,** 164–167 (1981), Figs 2 and 3 have been incorrectly numbered. Figure 2 is the graph of relative departures of average July temperatures from their means and Fig. 3 a time-series of estimated July temperatures.

MATTERS ARISING

Brown adipose tissue and diet-induced thermogenesis

IT has been proposed[1,2] that 'diet-induced thermogenesis' in brown adipose tissue enables animals to dissipate excessive intake of energy and so avoid obesity. The hypothesis is claimed to justify the view that thermogenesis is more important in regulation of energy balance than control of food intake[3], and the report by Rothwell and Stock[1] has been presented by the media (a BBC Television programme *Horizon*, 10 December 1979) as embodying an important discovery in relation to human obesity. We consider their data and arguments inadequate to support the hypothesis for the following reasons:

(1) The first paper hinges on an experiment in which one group of six rats was induced to overfeed by 'cafeteria feeding' (that is, providing a variety of tasty foods)[4,5] for 3 weeks, with a second group as controls. The amount of energy the cafeteria-fed rats stored as fat was much less than the increase in their intake of energy—this discrepancy supposedly demonstrated thermogenesis. We consider that it is hazardous to infer energy expenditure in this way (see Table 1 of ref. 1), particularly where it is the quantity of primary interest. (The paper includes in Table 2 some sample measurements of oxygen consumption as 'confirmation of lowered efficiency', but energy expenditure is too variable for short measurements to be useful in balance experiments, and the adjustment of the data for body weight complicates interpretation.) Accurate measurements of energy intake, expenditure and storage over a balance period are laborious and technically exacting; any systematic errors, however, cumulate in balance experiments, thus unless all three components of energy balance are measured there is no check on accuracy.

(2) Metabolizable energy intake (that is, the energy in food supplied minus the losses due to scattering and in excreta) cannot be reliably predicted from food composition tables or prior measurements: bomb calorimetry of the actual materials is essential. This must apply particularly to cafeteria feeding, where a variety of proprietary foods of high energy density are offered.

(3) Energy storage was determined by subtracting initial from final carcass energy values, determined by *post mortem* analysis; the method and its precision are not stated. The initial values were obtained from a group of six rats, killed at the start; their body composition was stated to be identical with that of the two groups used in the experiment, on the basis of *in vivo* tritium dilution measurements. We believe that estimation of body fat by *in vivo* body water measurement is inevitably inaccurate, and that this view is confirmed by Rothwell and Stock's data[6] on the tritium method.

(4) Rothwell and Stock[7-9] previously reported that cafeteria-fed rats showed similar resting oxygen consumption to controls and rapid increase of weight; that is, there was no evidence of thermogenesis (though it was reported to occur briefly after cafeteria feeding had ended). There are two possible explanations, other than experimental error, for these contrasting results. Rothwell and Stock[1,10] state that rats obtained from two suppliers, although of the same strain, differed in their capacity for thermogenesis and propensity to obesity. The only difference quoted in absolute terms was that energy expenditures during cafeteria feeding were 390 and 340 kJ per day. This difference, though statistically significant, is not large, and might be within the margin of error. If it is real, the selection of rats from one source rather than the other as the basis for a general theory would seem to require justification.

(5) We believe the age of the rats may provide the explanation. In Rothwell and Stock's earlier work[7-9] female rats weighed ~300 g and males 400 g and controls gained weight at 0.5–1.0 g per day, values typical of adult rats. The rats used in the later experiment[1] were described as 'adult' but their age was not stated; the control group were gaining weight at 5 g per day which is a rate typical of young, actively growing rats. Elsewhere Rothwell and Stock[2,10] describe as 'adult' rats which were 6 weeks, 10 weeks and 75 days old. Growth is costly of energy and its rate is sensitive to the supply of energy; thus it is well known that young, actively growing animals do not acquire fat readily. We believe a discrepancy between increased energy intake and fat storage, seen in young rats but not in adults, is more likely to reflect the young animals' capacity to use energy for growth than to demonstrate a new mechanism for energy dissipation. If the young animals' capacity for growth were to be regarded as a regulatory mechanism for energy balance, it would have to be made clear that it is not one available to adults. In adults, deposition of fat, although a relatively efficient process, entails dissipation as heat of about a third as much energy as is stored[11-13]; any discrepancy found in adults between excess intake and storage would have to be shown to exceed this before any possibility of a new mechanism arose.

(6) Since the work of Dawkins, Hull and others in the 1960s[15,21] it has been accepted that brown fat, controlled by noradrenergic sympathetic nerves, is an important heat-producing organ in the regulation of body temperature in infant animals of many species. Existing evidence indicates that the amount of brown fat in the body, and the heat produced in response to noradrenaline, decline with age[14-17]. The suggestion that, in animals possessing it, brown fat might be activated in response to overfeeding is an interesting one, although in young animals such a response would be expected to be antagonistic to growth. Miller and Payne[18] and Gurr *et al.*[19] have reported that diet-induced thermogenesis is conspicuous in young pigs; the pig is thought to possess no brown fat[20]. In our view, however, better evidence is required that diet-induced thermogenesis exists at all, before the question of its mechanism is worth pursuing. Extension of the suggestion to humans, and to adults, (and so relevance to human obesity) would seem doubly doubtful. The pair of IR thermograms[1] presented are certainly not adequate evidence: the effect seen after administration of ephedrine could as readily be ascribed to circulatory changes (G. A. Brown, RAF Institute of Aviation Medicine, personal communication).

Rothwell and Stock's hypothesis may still prove correct. Our point is that a novel theory which reverses previously accepted views needs substantial supportive evidence that can stand up to critical evaluation. We do not think such evidence has yet been presented.

G. R. HERVEY
G. TOBIN

Department of Physiology,
University of Leeds, Leeds LS2 9JT, UK

1. Rothwell, N. J. & Stock, M. J. *Nature* **281**, 31–35 (1979).
2. Brooks, S. L., Rothwell, N. J., Stock, M. J., Goodbody, A. E. & Trayhurn, P. *Nature* **286**, 274–276 (1980).
3. Perkins, M. N., Rothwell, N. J., Stock, M. J. & Stone, T. W. *J. Physiol.* (in the press).
4. Sclafani, A. & Springer, D. *Physiol. Behav.* **17**, 461–471 (1976).
5. Rolls, B. J. & Rowe, E. A. *J. Physiol.* **272**, 2P (1977).
6. Rothwell, N. J. & Stock, M. J. *Br. J. Nutr.* **41**, 625–628 (1979).
7. Rothwell, N. J. & Stock, M. J. *J. Physiol.* **291**, 59P (1979).
8. Stock, M. J. & Rothwell, N. J. in *Animal Models of Obesity* (ed. Festing, M. F. W.) 141–151 (Macmillan, London, 1979).
9. Rothwell, N. J. & Stock, M. J. *J. comp. physiol. Psychol.* **93**, 1024–1034 (1979).
10. Rothwell, N. J. & Stock, M. J. *Proc. nutr. Soc.* **39**, 20A (1980).
11. Blaxter, K. L. in *Obesity in Perspective* (ed. Bray, G. A.) 127–135 (DHEW Pub. No. (NIH) 75–708, 1975).
12. Pullar, J. D. & Webster, A. J. F. *Br. J. Nutr.* **37**, 355–363 (1977).
13. van Es, A. J. H. *Nutr. Metab.* **21**, 88–104 (1977).
14. Alexander, G. *Int. Rev. Physiol.* **20**, 43–155 (1979).
15. Smith, R. E. & Horwitz, B. A. *Physiol. Rev.* **49**, 330–425 (1969).
16. Heaton, J. M. *J. Anat.* **112**, 35–39 (1972).
17. Himms-Hagen, J. *Pharmac. Rev.* **19**, 367–461 (1967).
18. Miller, D. S. & Payne, P. R. *J. Nutr.* **78**, 255–262 (1962).
19. Gurr, M. I., Mawson, R., Rothwell, N. J. & Stock, M. J. *J. Physiol.* **291**, 60P (1979).
20. LeBlanc, J. & Mount, L. E. *Nature* **217**, 77–78 (1968).
21. Dawes, G. S. *Foetal and Neonatal Physiology* Ch. 15 (Year Book, Chicago, 1968).

ROTHWELL AND STOCK REPLY—We shall deal with points (1), (2) and (3) above together. Estimation of energy expenditure by the carcass balance method (that is, intake minus storage) is an accepted

0028–0836/81/080699—02$01.00

and reliable technique that has been in routine use in our laboratory for 15 years. We have introduced various improvements to increase accuracy and the coefficient of variation on carcass energy determination is less than 1% (ref. 1). A further improvement we have made is to match groups of rats at the start of the experiment (including those killed for determining initial carcass energy) on the basis of weight and body composition. In adult rats there is a poor correlation between body weight and energy content[2] and matching groups by weight alone or previous growth rate (as used by others, including Hervey and Tobin) can introduce very large errors. Obviously matching for body composition has to be done using *in vivo* measurements and, contrary to Hervey and Tobin's statement, our tritium method[3] gives values for body fat which do not differ from values determined by direct carcass analysis. However, after using the *in vivo* method for group matching, we then determine carcass composition of all groups (initial and final) by direct analysis, and this was clearly stated in our article[4].

We would share Hervey and Tobin's reservations about the use of food composition tables, but by carefully selecting appropriate foods and making meticulous measurements of food intake, metabolizable energy can be estimated conveniently and reliably. We have checked this against determinations by bomb calorimetry and the errors are small (always <5%)—this must be compared with the very large difference in energy intake (80%) on the cafeteria diet. All non-quantal measurements have errors and the skill of experimental design relies on producing differences that are not obscured by these errors. Thus, even if we take an extreme and unlikely 'worst case' example and assume the energy intake of cafeteria rats was overestimated by 10% and energy storage underestimated by 10%, we still find that cafeteria rats ate 62% more than controls. Of this additional intake, only 8% was stored and all the remainder was lost as heat—that is, as diet-induced thermogenesis (DIT).

Hervey and Tobin take us to task (point (4) above) for obtaining contrasting results in rats from different colonies, and yet they must be aware of the strain and intra-strain differences that are known to affect many aspects of metabolism (see for example, Miller[5]). Biologists cannot ignore these powerful genetic influences on metabolism, and selecting rats from different colonies is an attempt on our part to simulate in laboratory rats the wide individual variability seen in many animals, including man.

The rats used in our studies ranged from 200 to 400 g in weight (7–16 weeks of age). In terms of reproductive maturity these animals are past puberty and in anthropomorphic terms range from 'young adult' to almost 'middle-aged'. DIT can be induced at each age and to say (point (5) above) that age explains our results ignores the fact that rats of the same age can maintain weight on intakes that differ by up to 60% (ref. 6). If anything, our brown fat hypothesis might be used to explain why the incidence of obesity increases in later life when the amount and activity of brown fat is known to decline (refs 14–17 above). We cannot accept the argument that the energy cost of growth or fat deposition accounts for DIT for the following reasons. We have observed increased DIT when (i) the rate of growth is unaffected[6]; (ii) fat deposition is lower in one cafeteria group than another[7]; and (iii) when animals are hypophagic and losing body fat[4,8]. We have also calculated that the energy cost of increased fat deposition (when it occurs) accounts for less than 8% of the increased heat production[4] and even this assumes that all the fat was synthesized and none was derived from dietary sources.

We wonder why Hervey and Tobin accept the importance of brown fat in infant animals (point (6) above), for which there is no good quantitative evidence, and ignore the evidence provided by Foster and Frydman's meticulous study in cold-adapted adult rats[9]. Using their technique we provided similar quantitative evidence in cafeteria-fed rats[10] but they ignored this too, even though they are well aware of it. Our hypothesis is not intended to be exclusive and in our discussion of the pig experiments (ref. 20 above) we suggested alternative mechanisms of DIT in this animal which supposedly lacks brown fat, although it now appears that the young pig does possess some brown fat (D. L. Ingram, personal communication). The IR thermograms[4] are certainly not adequate to support a role for brown fat in adult man, but we doubt that 'circulatory changes' (personal communication, above) can explain how, after ephedrine, the temperature at the nape of the neck can rise above core temperature while other adjacent areas remain unaltered.

In terms of DIT, our work is not novel but merely confirmation of a phenomenon first observed at the turn of the century[11] and subsequently reported by many other workers (see ref. 4). Hervey and Tobin have ignored this earlier evidence and make no comment about the evidence for a thermogenic defect in genetic obesity (for example, *ob/ob* mice) perhaps because this implies a role for active thermogenesis in leanness. The only claim to novelty we can make is the suggestion that brown fat is the major effector of DIT and we see this as a positive development in an established line of inquiry and not a reversal of previously accepted views.

N. J. ROTHWELL
M. J. STOCK

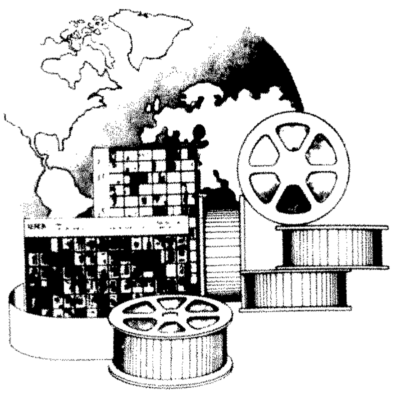
Department of Physiology,
St George's Hospital Medical School,
University of London, London SW17 0RE, UK

1. Lotfi, M., Macdonald, I. A. & Stock, M. J. *Br. J. Nutr.* **36,** 305–309 (1976).
2. Macdonald, I. A. thesis, Univ. London (1977).
3. Rothwell, M. J. & Stock, M. J. *Br. J. Nutr.* **41,** 625–628 (1979).
4. Rothwell, N. J. & Stock, M. J. *Nature* **281,** 3–35 (1979).
5. Miller, D. S. in *Animal Models of Obesity* (ed. Festing, M. F. W.) 131–140 (Macmillan, London, 1979).
6. Rothwell, N. J. & Stock, M. J. *Proc. Nutr. Soc.* **39,** 45 (1980).
7. Rothwell, N. J. & Stock, M. J. *Proc. Nutr. Soc.* **39,** 20 (1980).
8. Rothwell, N. J. & Stock, M. J. *J. comp. physiol. Psychol.* **93,** 1024–1034 (1979).
9. Foster, D. O. & Frydman, M. L. *Can. J. Physiol. Pharmac.* **56,** 110–122 (1979).
10. Rothwell, N. J. & Stock, M. J. *J. Physiol.* (in the press).
11. Neumann, R. O. *Archs Hyg.* **45,** 1–87 (1902).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

nature



is Available in MICROFORM

For Complete Information customers should write

University Microfilms International

Dept F A
300 North Zeeb Road
Ann Arbor, MI 48106
U S A

Dept F A
18 Bedford Row
London, WC1R 4EJ
England

TEXTBOOK SUPPLEMENT

Self-consciousness and the book for students

Students' books have at least one thing in common with fairy stories: they are written for the use of one section of the population by another. The authors of students' books are people who teach students and who pretend, sometimes accurately, to know what students need to know. Patronizing though this convention may seem, it appears to have survived even the fierce confrontation between students and teachers of the 1960s. The groups of students with ambitions then to help shape the university curriculum, often by the addition of "socially relevant" studies, seemed implicitly to accept that even though what to teach might be up for grabs, how to teach was for teachers to decide. Teachers have been more than willing to follow this prescription. In this convention, the contributors to the symposium of reviews of students' books which follows are teachers not students.

But does not this practice include a dangerous risk of circularity? If teachers write students' books, review them in journals such as *Nature* and also recommend them to their students, is there not a possibility that teachers' preconceptions of what students need will be self-justifying prophecies? At the level of the secondary school, there is ample evidence that this danger is not fanciful. Until the early 1960s, in Europe and North America, most science courses were marked by an almost ostentatious neglect of modern (or postwar) science. Now, some crusty teachers (mostly outside schools) argue, the pendulum has swung too far in the opposite direction. Students at schools learn about the relevance of DNA to human genetics before they know the first thing about the hydrogen bonds that hold these molecules together. Both the old and the new conventions about the content of science courses are indeed determined by teachers, acting as often as members of examination boards or members of curriculum study groups as in the classroom. Fortunately, the dangers are less serious at the level of the university.

The most obvious safeguard is the incoherence of the market in which the would-be authors of students' books must make their way. At some point in a teaching career, the impulse to see a successful course between hard covers, and launched on a wider world than that of the teacher's own university, seems to take hold. Publishers are willing accomplices, in normal times eager to make sure that their own portfolio of standard course books is comprehensive enough for them to offer almost everything to almost anybody (and also to keep their sales forces productively employed). Given only a modicum of independent judgement by teachers who are not themselves the authors of books, there is every reason to hope that the most intellectually successful course books will be those that sell most widely. Even when it seems that the simple influence of the market may be distorted by extraneous considerations — a textbook by a well-known author or from a distinguished institution — there is no serious reason to fear that students' teachers fail to recommend books honestly.

But is it proper that the important task of educating students should be determined in such a haphazard way? Might it not be better that university students, like those in schools, should be provided with books and other teaching materials that are the product of a self-conscious examination of students' needs. In short, should universities borrow from the schools and embark on deliberate curriculum development? In the past decade and more, this challenge has been cheerfully accepted in several places and in many fields of study. Groups of university teachers have spent their weekends hammering out a logical structure for some course of instruction, have gone back to their classrooms to try out the ideas on their students and have in the end subscribed their names to the title pages of one or more books. The result is not always the unsatisfactory compromise that might be expected. Especially when the objective is to devise a teaching course that crosses established disciplinary lines, the product of these endeavours may be stimulating and valuable. If, for example, the objective is to give the teaching of a subject a greater sense of history (a crying need in most fields, especially in mathematics), most practising subject teachers will acknowledge that they could profitably make use of help from others. To the extent that course-books are increasingly the product of several people who work together as a teaching team, some of the more obvious objectives of curriculum development are already assured. Yet formal curriculum development has inevitably hung fire outside the schools. Why should this be?

Some parts of the explanation are self-evident. Good teachers are rightly as suspicious of the all-embracing course-book as their better students are resentful of the intellectual straitjacket it represents. One essential part of a university education is to learn, mostly by hard experience, how to gather different nuggets of understanding from various sources, many of them books. The course-book is thus potentially an enemy of learning, a crutch and not a challenge. Second, those who teach in universities quite properly hold that it is their duty to teach their subjects in distinctive ways. It would be a great misfortune if this principle were eroded, for teaching is at its best a creative process.

These are sufficient reasons why a certain degree of anarchy in the production of students' books is to be welcomed. Another is that many books, originally devised as books for students, have become important components of the scientific literature. Feynman's three-volume *Lectures in Physics* has been for the past twenty years as much a challenge to his peers as to the undergraduates at the California Institute of Technology to whom the lectures were first delivered. In the same field but in a very different style, Max Born's *Atomic Physics* remains a classic of succinctness and a good read, even though the subject has been transformed in the past half century. R. H. Fowler's *Statistical Mechanics*, an enthralling graduate textbook, illustrates a different point; when amended to take more explicit account of the relationship between statistical mechanics and thermodynamics (as *Statistical Thermodynamics*, by Fowler and E. A. Guggenheim), the result was awkward and too didactic. Plainly, self-consciousness can be dangerous.

Self-consciousness is nevertheless not entirely to be outlawed. Indeed, there are many points in the undergraduate curriculum at which there is an urgent need of it. Why is it, for example, that chemists and physicists are still taught thermodynamics as if they were being taught quite different things? Why is the physics of engineering so different from physics as taught in physics departments? Why is the common ground between undergraduate organic chemistry and biochemistry so hard to find? Why are there nearly as many different kinds of "mathematics" as there are universities at which mathematics is taught? Why, throughout the undergraduate teaching of science, do teachers embark apparently independently on the design of experiments and projects, unable to profit from what has been done elsewhere? The dangers of self-consciousness notwithstanding, is it not time that university teachers wrestled more diligently with these problems?

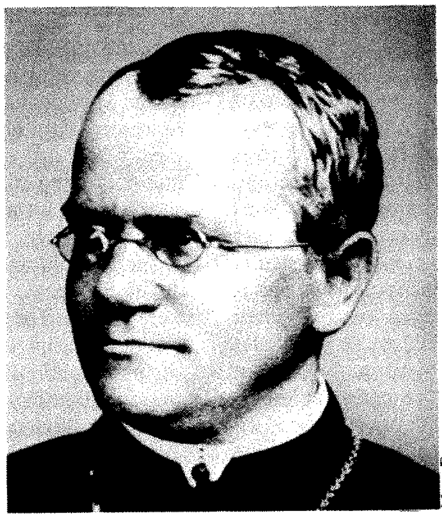
Mendel meets DNA: the teaching of genetics

D. J. Cove

Genetics. By C. J. Avers. (Van Nostrand: 1980.) *Modern Genetics.* By F. J. Ayala and J. A. Kiger Jr. (Benjamin/Cummings/Addison-Wesley: 1980.) *The Science of Genetics: An Introduction to Heredity.* 4th Edn. By G. W. Burns. (Macmillan, New York/Collier Macmillan, London: 1980.) *Principles of Genetics.* By J. W. Fristrom and P. T. Spieth. (Chiron/Blackwell Scientific: 1980.) *Biology of the Gene.* 3rd Edn. By L. Levine. (C. V. Mosby: 1980.) *Lecture Notes on Genetics.* By P. J. Russell. (Blackwell Scientific: 1980.) *Introduction to Modern Genetics.* By R. P. Wagner *et al.* (Wiley: 1980.) Further details of all books are given in the table accompanying the review.

THE publication in the past year of seven textbooks of genetics, all aimed to provide support for an introductory undergraduate genetics course, provides an opportunity not only for a comparative review of the books but also for a more general examination of how the subject should be taught at this level. The teaching of genetics poses problems both of strategy and of tactics. Strategically, a logical order of development is needed but not easy to devise. The two alternatives most commonly adopted are the classical or "Mendel before molecules" approach and the modern or "Let's start with DNA" approach. But in both of these there is the risk of compartmentalizing genetics and this risk is increased when quantitative, population and ecological genetics are considered as a third discrete topic generally tacked on behind.

The table accompanying this review indicates the gross order adopted in each of the seven texts, but the depressing point to emerge is that although some of the authors say they have opted for an historical approach (e.g. Avers; Burns) and some claim their approach to be modern and molecular (e.g. Levine; Russell), while others put forward their approach as neither historical nor modern but "born of pedagogical logic" (Fristrom and Spieth), there is rather little difference in the end product. The order of presentation is obviously different but there is little feel for the unity of the subject. Because formal, Mendelian genetics was established before biochemical genetics, we are not obliged to teach the two subjects separately; because quantitative and molecular genetics have developed largely independently of one another, the two subjects need not be kept apart for ever; and because the concept of dominance was introduced before the physiology of gene expression was understood, we are not compelled to continue to teach about dominance in an abstract and, to the student, often mysterious manner. There are of course risks in becoming a slave of the integrated approach and the sequential nature of learning inevitably imposes the need to teach some topics before, rather than alongside, others. Nor should it be forgotten that textbooks are not intended, at least in most cases, to be lecture notes for teachers. To be useful as reference works some compartmentation is desirable, but in many cases integration brings not only the advantage of reinforcing the unity of the subject of genetics but also allows some topics to be understood much more easily. The phenomenon of epistasis, for example, presents few problems to the student if introduced in the context of interaction between genes affecting the same biochemical or developmental pathway and this approach minimizes problems encountered by students in distinguishing epistasis from dominance, which are in molecular terms two quite distinct phenomena. It is interesting that Wagner *et al.*, a text which opts for the "Mendel before molecules" approach, treats epistasis in just this way, while Levine, a "Let's start with DNA" book, deals with epistasis without at the same time considering gene function. Curiously, Fristrom and Spieth, despite being a long text, fails to mention epistasis at all.


Gregor Mendel, father of modern genetics

Mary Evans

Before leaving teaching strategies, I want to return to overall order of presentation. There is no one order in which to teach elementary genetics which is superior or correct. Each teacher or textbook writer will favour his own individual approach. Some will wish to impose an evolutionary bias, others a molecular bias, some will favour a physiological and cellular emphasis, others will regard consideration of the whole organism to be more important. What is vital is to adopt a strategy which emphasizes how each field of genetics is dependent on each other field and how all are interrelated. In this respect I cannot recommend with enthusiasm any of the seven textbooks reviewed here, nor for that matter any other text published previously, although, at least in parts, Avers, Ayala and Kiger, and Wagner *et al.* make the most effort.

And so I come to tactics. It is fortunate that most students find genetics an interesting and stimulating subject since many undoubtedly find it difficult to master. Understanding the logic involved in connecting experimental observation to final deduction is a problem encountered across the whole of science but the connecting thread is especially difficult to follow in many genetical experiments. The logic involved is, however, crucial to genetics and is why the subject remains so powerful in the analysis of living systems. Teaching students to understand this logic so that they can themselves make deductions from data is one of the biggest problems facing the lecturer, who can be forgiven for being driven from time to time to the conclusion that geneticists can only be made in heaven. Here, a textbook can help the student if the relevant logic can be returned to again and again until it falls into place.

Unfortunately, the logical and rigorous way of treating a topic is not necessarily the most interesting and is certainly the most demanding. The device of giving the conclusions first and then showing how the experimental findings are consistent with these is the approach most often adopted, but this is a poor compromise, too easily deceiving the student into believing that he has understood the principles involved. The point can be illustrated by considering the use of deletion mapping to produce fine structure maps of genes. The most straightforward approach to teaching this topic and essentially that used by those of the textbooks under consideration which deal with it, is to take as a starting point an established deletion map, usually out of deference to history the rII genes of phage T4, and to show how point mutants can be located on this map within regions defined by the end points of the deletion mutants. Most students find this approach easy to follow yet, if asked to construct a deletion

---

D. J. Cove is Professor in the Department of Genetics at the University of Leeds.

Genetics textbooks — how they compare

	Avers	Ayala and Kiger	Burns	Fristrom and Spieth	Levine	Russell	Wagner *et al.*
Number of pages	659	844	608	657	542	258	573
Text potential (k words)[1]	360	430	320	410	300	190	360
% diagrams/figures/tables[2]	28	50	23	15	24	41	35
Problems	Yes	Yes	Yes	No[3]	Yes	No	Yes
Answers							
Given?	Yes	No[3]	Yes	No[3]	No	No	No
Quality of explanation	Fair	Good[3]	Poor	—	—	—	—
Order of presentation							
DNA before Mendel?	No	No	No	Yes	Yes	Yes	No
Meiosis before Mendel?	No	Yes and no[4]	No	Yes	Yes	Yes	No
Population, quantitative and ecological genetics last?	Yes	Yes	No	Yes	Yes	Yes	Yes
Price — hardback	£17.20	£9.95	£13.95	£11.50	NA[5]	NA	£11.15
	$22.95	$21.95	$19.95	$21.95	NA	NA	$26.20
Price — paperback	NA	NA	£6.95	NA	£12.75	£6.80	£6.45
	NA	NA	NA	NA	$17.95	$19.75	$14.15

[1] Text potential = number of pages of text × estimated maximum number of words which could be included on a full text page. [2] Estimated as the mean of 50 pages chosen at random. [3] Available separately in a "solutions manual". [4] Mitosis and meiosis introduced briefly in an introductory chapter but returned to later. [5] Not available with this binding.

map given raw data, many find the task extremely difficult. Here, a worked example might be helpful. Only Avers, Ayala and Kiger, and Fristrom and Spieth make some attempt at this but their treatments are not really adequate.

I have examined these texts more generally for their treatment of genetic logic, considering not only their approach to Mendelian genetics but also their treatment of experiments such as that of Crick and coworkers which elucidated the general nature of the genetic code and which remains an outstanding example of logical genetical deduction. On these criteria Ayala and Kiger perform best overall and genetic logic is one of Russell's stronger points. Fristrom and Spieth treat Mendelism well (but omit the Crick experiment). The other texts are more patchy; Avers acknowledges the "brilliant reasoning" of Jacob and Monod but does not attempt to duplicate it. Levine, although stressing in the preface the importance of presenting "wherever possible, the experimental procedures and data that have led to our present concepts in genetics", is the most descriptive, making little attempt to present even Mendelism in a logically rigorous way. All the texts, once again, display in their treatment of Mendelism, the extent to which they are constrained by the historical development of the subject. That Mendel was able to make his deductions concerning the mechanism of inheritance, despite the fact that the characters he used were expressed in diploid tissues and that the character differences showed complete dominance, makes him all the more to be admired. It is unnecessary however to expect students to cope with segregation, diploidy and dominance all at one bite. I favour introducing Mendelism by first considering segregation for characters expressed in haploid tissues and dealing next with a "diploid" character in which incomplete or codominance allows the immediate recognition of the 1:2:1 segregation in the F2 generation. This course is not without hazard, with students occasionally attempting to explain a pattern of inheritance in *Drosophila* by proposing that its tissues are haploid. The introduction of diploid segregation by way of incomplete dominance, however, has certainly proved satisfactory.

All the texts adopt the historical approach, in some cases appending incomplete dominance and haploid segregations as embarrassing exceptions, and in passing implying that dominance is an invariant relationship between two alleles rather than dependent on the level and aspect of the phenotype examined. Perhaps more seriously, the texts give the impression, initially, that each gene has only two alleles, one functional, the other mutant, an approach which results in the student having to backtrack when confronted with many topics introduced later. Ayala and Kiger once again perform best when these points are considered. Although describing complete dominance first, incomplete and codominance are introduced in the same paragraph, and the segregation ratios associated with them follow immediately after those obtained with complete dominance; multiple alleles are also introduced early on, but no emphasis is given to the artificiality of the concept of a unique wild-type allele.

Problem solving provides an excellent way in which the student of genetics can satisfy himself that he has understood the ideas which have been introduced to him. Surprisingly, since all seven texts are American, two contain no questions (but see table). Three contain questions but no answers, although in the case of at least Ayala and Kiger these are obtainable as a separate booklet. Burns contains both questions and answers but most of the answers are so brief as to be educationally worthless. Only Avers has, within the single book, questions together with answers which contain at least some explanation, although these are not as good as those provided by Ayala and Kiger in their solutions manual (which, rather curiously, is said in their preface to be "available to the instructor"). I see no point in including questions in a textbook unless answers, containing a clear and adequate explanation, are also readily available. Space is at a premium in all textbooks but in a genetic textbook set problems must have a high priority.

The authors of all seven books have obviously tried to make their books up to date. All contain at least some treatment of recombinant DNA technology and its implications but they do not all deal with intervening non-coding sequences in eukaryote DNA. Both Ayala and Kiger, and Fristrom and Spieth contain good if brief accounts of both these topics.

The texts vary very widely in their emphasis on figures and diagrams. Ayala and Kiger justifiably draw attention to the generous provision of illustrations in their book and must be given the prize not only for the clarity of their figures but also for general layout and design. Fristrom and Spieth is not only the most sparsely illustrated but has a layout which immediately looks dull. This is unfortunate for this textbook is certainly not dull in content. There is little to choose between the other five books with respect to design or their use of figures.

And so I come to where I am no doubt expected to name my best buy. No general textbook covering the whole of genetics can be expected to please everyone; inevitably, each will have strong and weak points. Although not perfect, Ayala and Kiger is, as I have indicated already, obviously superior on a number of counts. Despite its layout, I liked Fristrom and Spieth's treatment of many topics, including developmental genetics, and so I would make this my second choice. Of the rest, I am unsympathetic to the descriptive style of Levine, but others may be less critical; Russell, which is one of a "Lecture Notes on . . . ." series, is too superficial to be used as a reference book and does not provide as good value for money as the other books; the remaining three books all have good points — Burns, for example, emphasizes human genetics throughout, Avers integrates the genetics of discontinuous and continuous phenotypic variation, and an important point in favour of Wagner *et al.* is that I have already received favourable reports of it from my own students. However, I am disinclined to recommend any of the books except Ayala and Kiger, and Fristrom and Spieth, except perhaps for reference in libraries. □

Microbial genetics abbreviated

J.R.S. Fincham

Genetics of Microbes. By Brian W. Bainbridge. Pp.193. (Blackie/Halsted: 1980.) Hbk £16.95; pbk £7.95, $27.95.

THIS book deals with fungi as well as with bacteria and bacteriophage, and attempts to take the reader from the fundamentals up to the frontiers of research. It covers a reasonably comprehensive range of topics — some old, such as parasexual reproduction in fungi, and others rather new, including mitochondrial genetics of yeast, genetic analysis in streptomycetes, and *in vitro* manipulation and cloning of DNA in plasmid and phage vectors. The basic systems of gene exchange in bacteria, including transposons, are all described, albeit sketchily in some cases. Lambda phage gets considerable attention.

To cover all this in 170 pages (exclusive of bibliography), as Dr Bainbridge has attempted to do, is an ambitious undertaking, and I cannot say that he has been wholly successful. Conciseness is too often the enemy of clarity. Much information is presented in the form of complex diagrams, often based upon recent reviews and in themselves good and well chosen. But the captions are usually very short and often supported by only a telegraphic sentence or two in the text. Sometimes essential explanations or even facts are omitted in the rush to get on to the next topic. For example, after a brief account of gene conversion we come, a few pages later, to evidence from random meiotic products which is interpreted as favouring the idea of negative crossover interference. No hint is given that, in the light of tetrad analysis, data of this kind can be largely explained in terms of gene conversion. In the same chapter, an otherwise clear account of the Poisson distribution and its application to recombination theory is spoiled by a paragraph implying that a mean chiasma frequency of 0.5 corresponds to 50 per cent recombination (a factor of two missed here). In the chapter on the analysis of the yeast mitochondrial genome, the uninformed reader is not told that mitochondrial recombination is customarily detected and assayed in diploid hybrid clones and he will probably come away with the impression that it happens only in meiosis.

This, then, is not the ideal book, but it has its good points. The account of genetic engineering will be useful for beginners and the discussion of *Streptomyces* genetics provides information not readily available elsewhere except in the specialist literature. The chapter on DNA repair and mutation gives an accurate up-to-date account, albeit virtually confined to effects of UV. Several of the tables and figures provide good summaries of things one often wants to look up, and the bibliography is a useful guide to the primary sources. □

J.R.S. Fincham is Buchanan Professor in the Department of Genetics, University of Edinburgh.

Genes and behaviour

B. Burnet

Behavioral Genetics: A Primer. By R. Plomin, J. C. DeFries and G. E. McClearn. Pp.417. (W. H. Freeman: 1980.) Hbk £14.80, $24.95; pbk £8.10, $14.95.

BEHAVIOURAL genetics is the discipline linking the biological and behavioural sciences. It is concerned with the systematic study of the heritable bases of behaviour and the role of behaviour in evolution. The problem with teaching it is that intelligent discussion of the interesting issues depends upon a background knowledge of genetics, behaviour, statistics and evolution, and this is what the book provides. Following some historical background about Darwin and Galton, the authors discuss the theory of evolution in relation to behavioural genetics and sociobiology. There is a clear and concise introduction to genes and chromosomes and to the quantitative and population genetic approaches to the study of behaviour. Especially valuable is the careful explanation of how genetic variation in human behaviour is studied and how this yields information about the *environmentality* as well as the *heritability* of behavioural differences. They draw together the results of recent surveys which indicate that IQ differences are less heritable than earlier studies have suggested.

The authors take their material mostly from studies on human beings and small mammals except in their chapter "Mechanisms of Heredity and Behaviour", which includes bacteria, protozoa, round worms and insects. Those who, like myself, believe that invertebrates will for some time provide the arena for testing rival theories about the causal networks linking genes and behaviour may feel that advances using fruitflies and nematodes deserve more emphasis. The assertion on p.41 that "amphibians arose from the invertebrate lobe-finned fish" may discomfort vertebrate palaeontologists, and I note with interest a novel enzyme in Fig.10.1.

This is an excellent text, not only for the quality of the scientific explanations but because it deals critically with several controversial issues. It shows why behavioural genetics is intellectually exciting and practically relevant to the human situation. I recommend it for course work in the biological and behavioural sciences, and to the general scientific reader. □

B. Burnet is Reader in Animal Genetics at the University of Sheffield.

"Mini Luria"

D. H. Watson

Introduction to Modern Virology. 2nd Edn. By S.B. Primrose and N.J. Dimmock. Pp.251. (Blackwell Scientific/Halsted: 1980.) Pbk £7.50, $19.95.

IN this second edition of *Introduction to Modern Virology*, forming part of the Blackwell *Basic Microbiology* series, Dr S.B. Primrose has been joined by his colleague Dr N. J. Dimmock. The 251-page volume is amazingly comprehensive, forming a "mini Luria" in that it adopts the same general plan of Luria's *General Virology* (Wiley, 3rd Edn 1978). The treatment follows that given in a number of universities offering comprehensive virology courses as a component of undergraduate microbiology degrees.

An introductory chapter gives an historical survey of the development of ideas on the nature of viruses; a useful guide on basic techniques of animal virology is appended. Chapters on structure and nucleic acids of viruses are followed by six core chapters on the virus growth-cycle. Later chapters deal with such topics as virus pathogenicity, the immune system and interferon, vaccines and chemotherapy, and tumour viruses. A useful addition is a concluding chapter surveying virus families giving their "official" names, although the authors appear to be unaware of the recent U-turn of the relevant international committee which is dropping the unwieldy family name "herpetoviridae" in favour of "herpesviridae", after having strenuously resisted the arguments to do so in 1975!

The honours student will inevitably need to refer to more comprehensive volumes and a useful reference list is provided. Nevertheless, most teachers of virology would be more than delighted were their students to have a firm grasp of all the material in this book.

Inevitably, individual teachers will be slightly critical on one or two minor points. I was a little disappointed that the section on neutralization of viruses by antibody, reflecting the inclusion of Della Porta and Westaway in the reference section, appeared to suggest that the multiple critical site theory is fully substantiated. My preference would be to refer students to a more balanced view as represented in the reviews of Mandel. Others, perhaps, may be a little surprised at the omission of any reference to sub-unit vaccines in the relevant chapter. However a book of this size inevitably involves condensation of argument and selection of some topics at the expense of others, and there can be no serious claim of major omissions. The volume is likely to be welcomed by students and teachers alike. □

D.H. Watson is Professor of General Microbiology and Head of Department at the University of Leeds.

Spoiling the story

R. Whittenbury

Microbiology. 3rd Edn. By G. A. Wistreich and M. D. Lechtman. Pp.786 + appendices. (Glencoe, Encino, California/Collier Macmillan, London: 1980.) $20.95, £11.95.

GOOD plays, good novels, even good textbooks for undergraduates all display a common quality — a mastery of the subject, revealed in clarity of style, fluency of writing and in the skilful telling of the tale. *Microbiology,* I'm sorry to say, is not of this class. It is divided into six sections, three on basic principles of microbiology and three on related topics (immunology, medical microbiology and control of micro-organisms, and applied microbiology). Although much of the book is unexceptionable, it is badly flawed particularly in the area of basic principles.

A more personal criticism should be directed towards its "education props" style of presentation. All students (one likes to believe) relish a bit of an intellectual challenge. In this instance the challenge will, I fear, be boredom, rather than comprehension. Each chapter is prefaced by a brief description of the goodies to come and a list of major "take-home" messages, point by point, to be absorbed. In other words, the bones of the plot and the fate of the main characters are revealed before you've even had a chance to clap eyes on the chapter proper, itself subsectioned and highlighted. To reinforce this approach the chapters finish with a fairly lengthy sub-sectioned summary, again point by point — "Ribosomes are important in protein synthesis" — a list of review questions and a supplementary reading list (a bit too advanced for the readers the authors have in mind).

Excellent micrographs, black-and-white and colour photographs, and clear tables and diagrams (with one or two notable exceptions) all help to hammer home the message. One quibble I have is the pointless and irritating colour banding of the tables.

The more serious problems, emerging mainly in the basic-principle sections, stem from the remit the authors have set themselves — to write a text for students who have had some exposure to chemistry or biology, but not necessarily to both. These emerge as ill-written paragraphs and sub-sections, many errors of fact, contradictions, and over-simplification of concepts and complex issues. The last sometimes leads to incomprehensible statements or to truisms, some of which even raise spurious queries in the mind of the reader, for example "Phototrophs are noted for their photosynthetic activity". All phototrophs, of course, are photosynthetic organisms, by definition; the very form of this seemingly innocuous statement sows the seed of an idea that there may be some other and more important feature characterizing all phototrophs. Failure to integrate to ensure consistency, if not accuracy, adds to the confusion. For instance, what is written about autotrophs, photosynthesis, methane oxidizers and methane producers contains many examples of all these faults. Many of the errors (typographical included) are really inexcusable in a basic textbook.

Four final examples of the sort of thing you shouldn't find in a work of this nature are: a gross typographical error — "phoyodynnyhriv bacteria"; a fiction — that carbonates can be reduced via the cytochrome-mediated electron transport chain; a contradiction — that lignin is an exception to the rule of microbial infallibility (after giving earlier a list of bacterial and fungal species which break down lignin); a paragraph containing misinformation and needing no valedictory comment from me:

> The microorganisms in the rumen also provide other specific functions, such as the synthesis of amino acids and vitamins. Some microorganisms leave the rumen and are digested in other parts of the gastrointestinal system to serve as a major supply of proteins and vitamins for the ruminant. This is particularly important for the nutrition of ruminants because grasses are deficient in protein.

The authors have much remedial work to do before the fourth edition emerges and before the book, as a whole, can be recommended for use as a student text. Even if the text were faultless — and, potentially, it is a good example of its genre — my preference would still be for the Roger Stanier style (*General Microbiology* by R. Y. Stanier *et al.;* Prentice-Hall/Macmillan, London, 4th Edn 1977), an example of the sort of book I had in mind when writing the opening sentence. □

R. Whittenbury is Professor of Biological Sciences and Chairman of the Department of Biological Sciences at the University of Warwick.

Nearly Fungi perfecti

Peter E. Long

Introduction to Modern Mycology. By J.W. Deacon. Pp.197. (Blackwell Scientific/Halsted: 1980.) Flexi £6.50, $18.50.
Introduction to Fungi. 2nd Edn. By John Webster. Pp.669. (Cambridge University Press: 1980.) Hbk £30, $79.50; pbk £9.95, $21.95.

MYCOLOGY, traditionally a taxonomically orientated part of a botany honours syllabus, now finds a place in microbiology courses where greater emphasis is placed on fungal development and function. *Introduction to Modern Mycology* is a welcome, though belated, addition to Blackwell's *Basic Microbiology* series, reflecting the new approach. Very much a book for beginners, it is a clearly written, compact, explanatory account of fungal structure, growth and development, coupled to consideration of the role of fungi in nature. Taxonomy and morphology are confined to the introduction and there is a concluding chapter on ways of restricting fungal growth. The 14 chapters include an amazing number of topics, aided by lengthy figure captions that impart additional information. The author admits to neglect of some traditional areas; these include the mechanisms of ascospore and basidiospore discharge, while slime moulds receive scant attention. Such omissions are largely excusable, but I draw the line at his failure to make any mention of the lichens, which number roughly a quarter of all described fungal species. The illustrations are not so well chosen as the contents and sometimes appear crude, especially in

The Great Gray Owl: Phantom of the Northern Forest, reviewed in *Nature* **288**, 519, 1980, will be published in the UK, Europe and the Middle East in March of this year by Oxford University Press. Price will be £9.95. The book is available in the USA from The Smithsonian Institution or through bookshops.

comparison with the next volume. It is a pity that the selection of further reading, mostly other books, is so restricted since it will not satisfy the interest in fungi that the book will arouse. Dr Deacon's book is currently the best, short, undergraduate introduction to fungi on the market, albeit at the cost of some oversimplification.

The second, expanded, edition of *Introduction to Fungi* is, by contrast, a reference work, suitable for second- or third-year undergraduates and of great value to all mycologists. It retains the format, style and superb standard of illustration found in its predecessor. Thus it creates the impression of a traditional mycology textbook. This is deceptive, as within its taxonomic framework lies an enormous amount of information on fungal structure, cytology, genetics, physiology and behaviour, as well as other general topics. All the subject matter is presented with precision and detail backed by innumerable references to original sources. The pre-existing chapters have been thoroughly revised, the treatment of slime moulds has been greatly extended and there are two new and welcome chapters on the general attributes of the Eumycota and on the Fungi Imperfecti. More attention than before has been paid to the production of the book, resulting in better placed and printed illustrations, a clearer hierarchy of headings, and more obvious stressing of terms through use of bold type instead of italics. In all, it is a considerable improvement on the first edition.

However, instructors who have to decide between recommending this book and the latest edition of Alexpoulos' *Introductory Mycology* (Wiley, 2nd Edn 1979) are placed in an unenviable position. Professor Webster undoubtedly provides a greater body of information about the fungi, in a work that is as helpful in the laboratory as it is in a study. Unfortunately the book is neither so easy to read nor apparently so topical as its American counterpart. This is largely a matter of presentation. *Introduction to Fungi* has only seven long chapters, necessitating numerous subheadings. In some cases it is not clear where a particular section ends. For example, "Breeding for resistance" (p.509) seems to continue into a discussion on rust life-cycles. Students will also encounter difficulties, at times, in sifting essentials from the mass of ancillary information as key points are not always well emphasized.

Any book of this length and quality, for which the author also takes on the onerous task of illustrator, carries the risk of falling just short of perfection. The present volume comes very near to escaping this fate. Despite its drawbacks for novices, it will undoubtedly and deservedly enjoy a wide market and will be regarded by many as a classic.

Peter E. Long is a Lecturer in Mycology in the Botany Department, University of Leicester.

Plant taxonomy surveyed

J.G. Vaughan

Plant Taxonomy and Biosystematics. By Clive A. Stace. Pp.279. (Edward Arnold/University Park Press: 1980.) Flexi £8.95, $29.50.

PLANT taxonomy in the modern context is very much a multidisciplinary subject and encompasses various types of botanical investigations and techniques. According to the author, this book is a broad survey of the subject and is aimed primarily at university undergraduate students and others in similar establishments. The information presented is largely limited to green plants, with the emphasis on angiosperms; fungi and lichens are excluded.

The first section deals with the scope of taxonomy and the types and development of systems of classification, the second with the sources (structure, chemistry, chromosomes, breeding systems, plant geography, ecology) of taxonomic information. The book is rounded off with discussion of the process and mechanics of classification. The treatment of the subject is thus fairly comprehensive and the text is well supported by references. Bearing in mind that the book is intended for undergraduates, I do have some minor criticisms— Fig. 3.7, showing 31 types of stomatal complex, could be daunting to such students, and the method of double-diffusion serology illustrated in Fig. 4.3 is not typical of that used by chemotaxonomists. Also, the distinction between lemon and orange on skin colour (p.225) would not be significant in West Africa, and possibly some other areas.

Stace's book is an important contribution to the literature of plant taxonomy in that it deals with all the classical and modern facets of the subject in a concise and readable fashion. Undergraduates, after a basic training in plant science, will find it most useful. Indeed, it is a valuable work for any scientist or institution seriously interested in plant variation.

J.G. Vaughan is Reader in Plant Taxonomy at Queen Elizabeth College, University of London.

Transport for botanists

James F. Sutcliffe

Transport in Plants. By U. Lüttge and N. Higinbotham. Pp.468. (Springer-Verlag: 1979.) DM58, $30.50.

RECENT research on the transport of materials in plants has been reviewed extensively in the first three volumes of a new series of the *Encyclopedia of Plant Physiology*, published by Springer-Verlag in 1975–1976. The same publishers have now issued a further book on the subject, based on Professor Lüttge's *Stofftransport der Pflanzen* (1973). It is a concise, authoritative and up-to-date account of research in a field to which the authors themselves have made important contributions.

Inevitably, in a book of moderate size covering a wide subject, some topics receive less detailed treatment than others. In general, the material reflects the research interests of the authors and so there is extensive coverage of membrane transport and its relationship to metabolism. On the other hand, phloem transport is disposed of in about ten pages and transport of water receives hardly any attention.

The first part of the book is devoted mainly to a straightforward account of various physical phenomena associated with transport processes, including osmosis, electrochemical potentials and electrogenic pumps. The Nernst equation, Goldman voltage equation and Onsager coefficients are explained. The presentation is clear and the mathematics involved should not be beyond the comprehension of a physiologically-minded biologist.

Part II is concerned with transport into cells and begins with a summary of current views on the structure of cell walls as a prelude to consideration of the concept of free space. This is followed by an account of the structure of membranes and various ideas on membrane-transport. The authors then turn to the complexities of the multicompartmented, vacuolated plant cell. Some background knowledge of cytoplasm is assumed. There is a useful appraisal of the value and limitations of the kinetic approach to the analysis of ion fluxes and an impartial assessment of the concept of dual mechanisms.

In Part III the control of transport and its relationship to metabolic processes are considered. Particular attention is paid to the role of photosynthesis in regulating ion-fluxes in algal cells and in isolated chloroplasts, and there is an interesting chapter on the influence of phytochrome and growth-regulating substances on membrane transport.

The final section of the book deals with inter-cellular and long-distance transport within plants, and covers such topics as the movement of ions across the root cortex, functioning of glands, ion-relations of stomata and transport in sieve tubes. Non-botanists will find this part of the book difficult to follow if they have no previous knowledge of plant structure and some simple line-diagrams illustrating the salient features would have been helpful.

The book is attractively produced and, in all, a useful volume. I shall recommend it enthusiastically to advanced undergraduates embarking on courses in plant physiology and to their teachers. □

James F. Sutcliffe is Professor of Plant Physiology at the University of Sussex.

World of botany

Peter D. Moore

Vegetation of the Earth (and Ecological Systems of the Geo-biosphere). 2nd Edn. By Heinrich Walter. Pp.274. (Springer-Verlag: 1979.) DM 29.40, $15.40. *Plant Geography.* 2nd Edn. By Martin C. Kellman. Pp.181. (Methuen: 1980.) Hbk £10; pbk £4.95, $10.95.

I AM sure that all ecologists and biogeographers will welcome a second edition of Heinrich Walter's *Vegetation of the Earth.* Perhaps the two most valuable features of the first English edition were its simplicity and clarity of terminology and arrangement, and its liberal use of physiological data, so well selected and integrated with climatic and distributional information. Although this new edition is slightly expanded, these two features do not appear to have been substantially strengthened. Clarity has, in my opinion, not been enhanced by the introduction of a new system of jargon. We now have eubiomes, zonobiomes, orobiomes and pedobiomes, where once we had deserts, savannahs and forests. The salt desert of Utah is now a eubiome of the pedohalobiome of zonobiome VII. Such jargon rivals terms like "dedesertification"! On the physiological side, the book remains rich in useful examples, but some omissions are conspicuous. I find it surprising that space is not found for a discussion of C_4 photosynthesis and its intriguing biogeographical implications; crassulacean acid metabolism is given only a few lines.

On the positive side, the book's expansion now permits coverage of geographic areas neglected in the first edition. This will be particularly welcomed by Australians and New Zealanders, who now receive more than passing mention — an unusual feature in any biogeography book. Overall, Heinrich Walter's book remains the best, condensed account of world vegetation on the market.

Martin Kellman's second edition of *Plant Geography* has been extensively rewritten and the material substantially reorganized. Three themes permeate the book: plant population biology, environmental history and evolution. It is a difficult task to bring together such extensive subjects at an undergraduate level and to attempt their integration. This is especially so if one feels it is necessary to start from a very basic level of biological background; within half a page he has defined and discussed mitosis, meiosis, polyploidy and introgressive hybridization. Fortunately, more resources are allocated to survivorship curves, seed demography and r and K selection. Species interactions and responses to the environment are considered, but here the accounts suffer from a lack of physiological information. The development of the interaction theme through coexistence to plant community analysis is logical, but there follows an unconformity, with chapters on techniques and then the human factor being slipped in as brief appendices.

Kellman aims in this book to stimulate in the minds of his readers questions relating to the evolution of plants and plant distributions. His careful selection of interesting topics should enable him to achieve this end, and his arrangement of material has certainly resulted in the second edition being a considerable improvement on the first. □

Peter D. Moore is Senior Lecturer in the Department of Plant Sciences, King's College, University of London.

How to stimulate students of plant physiology

L. J. Audus

Plant Hormones and Plant Development. By W. P. Jacobs. Pp.339. (Cambridge University Press: 1979.) Hbk £21.50, $16.95; pbk £7.25. *Biochemistry and Physiology of Plant Hormones.* By T. C. Moore. Pp.274. (Springer-Verlag: 1979.) DM 49, $27.

THE advent of extremely sensitive and chemospecific assay techniques has revolutionized plant hormone studies in the past two decades, and has brought into question many time-honoured theories of the hormone control of plant growth and development. What, then, do I look for in books such as these, intended as they are for undergraduate students of plant biology? Apart from a critical and balanced assessment of the present state of our understanding, I would hope to find that kind of stimulating presentation and analysis that would fire the student to go further into this fundamental science.

Of the two books under review, only Jacobs's comes anywhere near to meeting these requirements. He believes, and rightly so, that it is essential to emphasize "the development of research ideas" and this he has done quite admirably by taking carefully selected phenomena with which he himself has been most actively concerned (vascular development, hormone transport, abscission, flowering, for example), and making exhaustive, rigorous, critical yet constructive assessments based on his well-known PESIGS rules for evaluating the hormone evidence. Such treatment makes exciting reading which is marred only by occasional lapses into somewhat facetious and unnecessarily caustic criticism of other workers' deficiencies. It is, on his own admission, a very personal treatment, as witness for example the inclusion of data from his own researches in more than half of the plentiful illustrations in the text. Biochemical matters are virtually absent and the highly selective treatment leaves some important phenomena untouched (for example, geotropism) and the cover of the various hormone actions much restricted (for gibberellins only the action in flowering is adequately presented). Regrettably, ethylene is regarded as not being a hormone.

On the other hand Moore's book is predominantly biochemical, even molecular-biological. While Jacobs presents his chapters under phenomena, Moore does so under individual hormones, thus largely omitting the important matter of hormone interactions. Developmental responses are briefly covered but few (for example, dormancy) are treated satisfactorily. Detailed consideration of phytochrome, although systematic and complete, seems to be out of place in this book, especially since the role of hormones in flowering is so over-condensed and incomplete as to be dangerously misleading to the student. On the whole, hormone biochemistry is well treated; for example there is, among others, a nicely balanced appraisal of modern ideas of auxin mechanism.

While Jacobs is rigorously honest in the matter of priorities, Moore is occasionally slack; in fact his whole approach to references is puzzling. References to key researches are often not given in the lists at the end of each chapter, although these lists are much swollen by references to works which are not mentioned anywhere in the text. The student will not find this helpful.

Of the two, then, Jacobs's book would be my choice for recommendation to students, although it still falls short, in my opinion, of the earlier book by K. V. Thimann *(Hormone Action in the Whole Life of Plants;* University of Massachusetts Press, 1977) which fulfils all Jacobs's criteria concerning the development of ideas. It is just as rigorous, objective and stimulating but is better in that it maintains just about the right balance between molecular biology and developmental physiology. □

L. J. Audus is Emeritus Professor of Botany at Bedford College, University of London.

Biological "wheres"

Paul A. Colinvaux

Biogeography: An Ecological and Evolutionary Approach. 3rd Edn. By C. B. Cox and P. D. Moore. Pp.234. (Blackwell Scientific/Halsted: 1980.) £7.50, $19.95. *Biogeography.* By E. C. Pielou. Pp.351. (Wiley-Interscience: 1980.) £13.30, $31.30.

THESE are two books about the new biogeography, one a survey for undergraduates with little background, and the other a provocative series of essays touching on the state of the art.

Biogeography was the founder subject of both evolution and ecology. In Alfred Russel Wallace's *The Geographic Distribution of Animals* of 1876, evolution was triumphant and Sclater's six zoogeographic realms, each full of animals partly evolved in isolation, became evidence for evolutionary theory. The circumpolar zones of Allen, which recognized ecological realities — boreal forest, temperate forests, rain forests in rings of latitude on every continent regardless of ancestry — were ridiculed by Wallace as the study of mere adaptation. The Sclater-Wallace classification lasted until fairly recently, through to the plant geography of Good. In the intervening years, ecologists almost avoided the word "biogeography" and talked instead of life in the great formations, roughly corresponding to Allen's old zones, later to be known by Clements' unnecessary and regrettable term, biome.

But with ecology taken away, biogeography could be little more than the study of relationships towards the goal of a better phylogeny. True, a speculative fringe found excitement in continental drift and land bridges, but they made few geological converts.

For half a century, evolutionary biogeographers on an Earth of fixed continents proceeded to fill in details. At the last, Darlington in *Zoogeography* (Wiley, 1957) told us of endless radiations from the heartland of Africa (the Old World Tropics), perhaps feeling the ghostly presence of a rejected Gondwanaland.

Then came renewal. Remanent magnetism and plate tectonics gave us back Wegener's drifting continents; Pangaea and Gondwanaland lived. But even more vibrant was the incursion of ecology, as theorists of the Hutchinsonian school taught us to look at colonization, dispersal and speciation as ecological processes on a geographical scale. MacArthur and Wilson gave us *The Theory of Island Biogeography* (Princeton University Press, 1967).

Cox and Moore, in their third edition, press on with their successful look at this new biogeography through ecologists' eyes. Not only does the species equilibrium on islands stand with drifting continents as a central theme, but prime ecological concepts like Lindeman efficiency, niche and resource-partitioning find their places also. A possible worry is that ecological concepts that really require a mastery of competition theory have to be handled with the logistic hypothesis and its derivatives left out — competition is taken for granted rather than debated. Yet Cox and Moore make a brave show of giving modern ecology to students with no training in it, all the same.

Pielou's book is a survey by a master ecologist to be used by, as she puts it, "Senior undergraduates, and interested biologists and earth scientists of all kinds". She says that she is "probably rash" to attempt such a thing, but she was not. She has done what she said she would so well that no scientist of a related discipline can now be excused ignorance of the possibilities of the new biogeography.

Each of the ten chapters of Pielou's work will stand by itself as an essay, so you can really begin to read the book where you like. I recommend starting with the third, "Evolution, Phylogeny, and Geographic Spread". This sounds conventional enough; the title could be in a book from the 1930s. But you will instantly be drawn into a splendid analysis of modern systematics controversy that I just could not stop reading, even though in bed with 'flu at the time. Have you wondered what was behind the guerrilla war between Halstead and the Natural History Museum (*Nature* **288**, 208; 1980)? — read Pielou on cladistics, phylogeny and the evidence of biogeography.

She is a little hard on geologists who would not accept the biogeographic evidence for continental drift before the advent of plate tectonics — the honest truth was that it did not seem parsimonious to move a continent in order to accommodate a fossil. She puts island biogeography somewhere in the middle of the book, as a subject appealing to sectional interests, then follows that by a chapter on "Geographical Ecology". From then on the book wields the Hutchinsonian ideas of the evolutionary play in the ecological theatre, until we bow out with thoughts on the fundamental mechanisms of polyploidy and gene flow in the last chapter. It is a book both to read and refer to, highly modern but based on profound understanding of the history of an ancient subject. Many a graduate student will have both thoughts and work shaped by it. □

Paul A. Colinvaux is a Professor in the Department of Zoology at The Ohio State University.

Understanding ecosystems

I. J. Linn

Ecology: A Textbook. By H. Remmert. Pp.289. (Springer-Verlag: 1980.) Pbk DM 37, $21.90. *Ecology.* 2nd Edn. By R. E. Ricklefs. Pp.966. (Chiron/Nelson: 1979.) Hbk $22.95, £20; pbk £12.50. *Principles of Ecology.* By R. Brewer. (Holt, Rinehart & Winston/Holt Saunders: 1979.) $22.95, £10.50.

WHEN I was a student, many years ago, my early ecological diet was firmly founded in L. A. Borradaile's *Animal and its Environment* (Frowde, 1923) and A. S. Pearse's *Animal Ecology* (McGraw-Hill, 1926). Both of these books devoted much space to a detailed examination of the fascinating ways in which animals are adapted to survive in a wide range of habitats — the sort of thing which nowadays *Scientific American* does rather well. It is pleasing to note that in the new edition of Robert Ricklefs' standard work, despite some changes, he continues to open his argument with a solidly evolutionary approach to the adaptation of organisms to their physical and biotic environment. Of course he goes far beyond that — much has happened in half a century — and for me Ricklefs' *Ecology* is still the basic ecological reference book that no biologist (ecologist or not) can afford to be without. The jargon is all there; if you are not sure what is meant by "character displacement", or of the precise difference between "exploitation" and "interference" competition, the lucid explanations in text and glossary will help you out. For the numerate, there is liberal use of the necessary descriptive equations; and for those, like myself, whose command of matters mathematical is somewhat patchy, there is help in succinct English.

There are faults in the book, of course, but by and large they are minor. The major weakness is that, while the author by no means ignores work and examples from outside the Americas, his concentration on matters American (no doubt deliberate) amounts to a regrettable bias, not only for his readership outside America, but also for his domestic readers, who would benefit from broader horizons. As I scan my own bookshelves, I see that the great majority of books on ecology in English produced in the last 20 years or so are American, so it is a particular pleasure to welcome Hermann Remmert's *Ecology: A Textbook,* in which many European and some African examples are found alongside the American ones.

Remmert's book is only a third of the size of Ricklefs', so the treatment of the subject is bound to differ. Remmert has decided to concentrate on autecology, population ecology and ecosystems, and this approach is largely successful. The style of the book is also very different. You will find no jargon here, and very few equations, only plain and most readable English (the translation from the German by Dr Marguerite Biederman-Thorson is excellent). The result is a first-rate exposition of many of the basic concepts of ecology which can be read with profit by biologist and non-biologist alike. Professor Remmert is very concerned that the magnitude of the task of understanding complex ecosystems should not be underestimated, and he is sceptical of the value of mathematical models. He also has some harsh words for what he calls "ecological quacks" who offer advice on how to improve mankind's lot, and he points out that the need is to maintain in good condition the environment to which we are so well adapted, and that improvement in this context is a meaningless concept.

For this reason, presumably, he has included no section on applied ecology, unlike Richard Brewer, whose *Principles of Ecology* is otherwise similar in size and coverage, though with the expected American bias. Brewer, from his comments on Spaceship Earth, is equally concerned to conserve essentials, and his sections on the ways in which the ecologist can go about offering sensible advice seem to me not unreasonable, though I doubt if they would satisfy Professor Remmert. ☐

I. J. Linn is a Senior Lecturer in the Department of Biological Sciences of the University of Exeter.

Aquatic biology, past and future

C. M. Yonge

Fundamentals of Aquatic Ecosystems. Edited by R. S. K. Barnes and K. H. Mann. Pp.229. (Blackwell Scientific: 1980.) Flexi £7.80, $22.50. *The Ecology of Streams and Rivers. Studies in Biology,* 122. By Colin R. Townsend. Pp.68. (Edward Arnold: 1980.) Flexi £2.10.

WHEN I was on the staff of the Plymouth Laboratory the basic methods for the quantitative determination of phosphates and nitrates in the sea were being elaborated, while, at the same time, a descriptive shore ecology was being supported by experiments revealing such unsurprising facts as that organisms highest on the shore are most resistant to the effects of desiccation. Certain of us were making tentative advances in comparative physiology.

These two admirably presented and illustrated books are almost entirely concerned with the results of half-a-century of subsequent progress. The properties of the ecosystem emerge as no less significant than those of the organism or even populations of organisms. Energy flow is revealed as all-important.

The more comprehensive *Fundamentals of Aquatic Ecosystems* is the ably edited work of ten authors, each highly qualified to deal with some aspect of the subject which includes the increasing human exploitation, by aquaculture, of the products of marine and fresh waters. No better introductory survey of these highly significant aspects of modern biology could be desired, one certain to fire the imagination of senior pupils and first-year undergraduates. How rewarding, surely, to teach students already so well informed.

Relevance to the future arises out of comment in the book on the recent discovery near the Galapagos Islands of unique animal communities around submarine volcanic vents at depths below the penetration of light. Consequent description of the sulphur cycle, on which this life depends, involves mention of the industrial production of sulphur dioxide and its influence on the basic chemistry of the Earth. Such effects on the biosphere emerge as among the major hazards of coming decades.

The smaller *Ecology of Streams and Rivers* serves almost as a supplement, concerned as it is with the maintenance of populations within running waters with the constant danger of being carried seaward to destruction. It is an excellent addition to the *Studies in Biology* produced by the Institute of Biology. ☐

Sir Maurice Yonge is an Honorary Fellow in Zoology at the University of Edinburgh.

HARVARD
Paperbacks

Martin L Cody &
Jared M Diamond, editors
Ecology & Evolution of Communities
£8.75 (cloth £24.50)
"An authoritative account of the current state of knowledge and indications of future directions in this branch of ecology." — *Science*

Sarah Blaffer Hrdy
The Langurs of Abu
Female & Male Strategies of Reproduction
£4.80 (cloth £12.25)
"Seldom . . have new facts, information from the literature , and theory been so well integrated, the data so lucidly presented, and the text so skillfully written." — *American Scientist*

W John Smith
The Behavior of Communicating
An Ethological Approach
£4.80 (cloth £17.50)
". . an important book, not merely a review, but original, provocative and stamped with the author's highly personal approach to the subject. It is a 'must' for the specialist."
— *Journal of Animal Ecology*

Geerat J Vermeij
Biogeography & Adaptation
£5.40 (cloth £17.50)
"This Darwinian book — there is hardly higher praise — is a step toward understanding, which will probably challenge our technical methods more than the broad insights of that clever naturalist on the *Beagle* did." — *Scientific American*

Edward O Wilson
Sociobiology
The Abridged Edition
£7.00 (cloth £12.95)
"A new synthesis, of wide perspective and great authority . . . Wilson's plain uncluttered prose is a treat to read . . ." — V.C. Wynne-Edwards, *Nature*

HARVARD
University Press
126 Buckingham Palace Road,
London SW1
Circle No.29 on Reader Enquiry Card.

Concepts of animal behaviour

Aubrey Manning

Introduction to Ethology. By K. Immelmann. Pp.237. (Plenum: 1980.) $22.50, £14.18.

IMMELMANN defines ethology as "the study of the biology of behaviour". This covers almost everything, but whilst I am no longer convinced that ethology as a separate field exists, it remains easy to recognize the influence of the ethological *approach* in all branches of behavioural science. Certainly, we need textbooks which introduce ethological concepts, particularly as so many students are now plunged directly into elaborate sociobiological arguments and tend to give short shrift to the study of an individual's behavioural organization and development, which has always been ethology's strong point. The translation of Immelmann's book is going to be useful for it covers a lot of ground.

My heart sank at first when I saw, yet again, illustrations of Tinbergen's herring-gull bill patterns and stickleback models. Yet there are also plenty of fresh examples and it is healthy to have many from the German literature. Anyway there is nothing to say that old examples (or old concepts) are not still the best, provided that, where necessary, they are re-interpreted in the light of modern thinking. I do feel Immelmann falls short here in some cases, for example when describing the effects of the hawk-goose silhouette on game birds or when discussing terms like "action-specific energy" as causal factors in behaviour.

From a student's point of view the book is clearly set out, with sensible use of sub-headings. Systematization is taken too far in places, with Immelmann's excessive use of formal terms to classify his material. Some of these, such as "ethoendocrinology" and "acquired releasing mechanism (ARM)", are unlikely to be encountered again outside the pages of this book.

I would have liked more examples to be spelt out in detail, but Immelmann has a large field to cover and in general he summarizes very clearly. The chapter on social behaviour is particularly valuable as a modern re-statement of the ethological approach with due emphasis on sexual behaviour and care of the young. In summary, a good book to work through with students on an introductory animal behaviour course. □

Aubrey Manning is Professor of Natural History at the University of Edinburgh, where he works on and teaches animal behaviour. He is author of An Introduction to Animal Behaviour *(Edward Arnold, 3rd Edn, 1979).*

Primates in the lecture theatre

W.C. McGrew

Social Behaviour in Primates. By Neil Chalmers. Pp.256. (Edward Arnold/University Park Press: 1980.) Flexi £6.50, $19.95.

THE rise of behavioural studies of non-human primates over the past two decades has stimulated a concomitant growth of university courses, especially in North America. Because of their cross-disciplinary nature, such courses are taught in departments of anthropology, biology, and psychology, thus creating a large market for textbooks. The first wave of such books appeared in the early 1970s, Jolly's *Evolution of Primate Behavior* (Macmillan, 1972) and Rowell's *Social Behaviour of Monkeys* (Penguin, 1972), for example, but these are obsolescent.

Now comes a new entry into the area, written by an English primatologist experienced in African field-studies. The substance of the book is in six chapters — on groups, social structure, infant development, sex, dominance and adaptiveness. It focuses on fieldwork but calls in laboratory studies when appropriate, for example on hormonal effects. The approach is selective rather than encyclopaedic, and one of the author's strengths is his skill in choosing apt examples. The most serious drawback is that consideration of the function of behaviour is postponed until the final quarter of the text. What precedes is solid discussion of the causation of behaviour, but until the end there are no hooks upon which the reader can hang the obvious questions which this discussion stimulates. For example, why should primates live in groups, unlike many other mammals? Why do some, but not all, primates breed seasonally?

Still, these are but quibbles, as Chalmers has produced an admirably lucid textbook which deserves to be widely used. □

W.C. McGrew is a Lecturer in Psychology at the University of Stirling and a Visiting Faculty Member of the University of North Carolina at Charlotte.

Twelve years after the appearance of the second edition, Cambridge University Press have published a new edition of *Excursion Flora of the British Isles* by A.R. Clapham, T.G. Tutin and E.F. Warburg. The new edition is available with a plastic cover and costs £12.50.

More neuroscience

Trevor W. Robbins

Behavioral Neuroscience: An Introduction. By C. W. Cotman and J. L. McGaugh. Pp.838. (Academic: 1980.) $21.95, £13.80. *Physiological Psychology.* By T. S. Brown and P. M. Wallace. Pp.634. (Academic: 1980.) $20.95, £11.20. *Fundamentals of Human Neuropsychology.* By B. Kolb and I. Q. Whishaw. Pp.501. (W. H. Freeman: 1980.) $17.95, £8.95.

"How could anyone know enough to write a whole textbook?" Brown and Wallace thus modestly quote one of their own students in the preface to *Physiological Psychology.* Indeed, the sometimes dazzling technical advances in neuroscience have led to a proliferation in knowledge which defies the capacity of individual writers to achieve the degree of synthesis or balance required of a text. Brown and Wallace's book, however, is but one in a long line of such attempts and so it seems reasonable to consider some criteria by which textbooks in this area can be assessed.

One of my own criteria is to ask whether it is possible to guess from the relative qualities of the various sections, the specific research interests of the authors. I concluded (partly correctly) from *Physiological Psychology* that one or other of its authors had special expertise in the chemical senses, sexual motivation or circadian rhythms, since these sections seemed the most detailed and best-explained.

Most of the other topics covered fit into the familiar organization of chapters on introductory neuroanatomy and neurophysiology, sensory systems, motor systems, neural plasticity, learning and memory, and "higher processes". These subjects are generally covered adequately, but suffer from temptations seldom avoided by textbook writers in physiological psychology.

The first of these is to act essentially as an uncritical, passive reporter of what are considered to be main issues or experiments. For instance, Brown and Wallace's discussion of hunger and satiety flits uneasily between the vagaries of "central" and "peripheral" theories, for example, "But as we shall see later in this section, the pendulum is swinging back again, and is now somewhere in the middle" (p.291). Now this may be an accurate assessment of the state of that particular art but this lack of intellectual guidance would not instil me, nor I suspect, an undergraduate, with much confidence.

A second, related temptation is to avoid such shows of indecision by skirting any real discussion of the questions posed, or of the nature of the underlying processes. Thus, an otherwise tidy account of "brain stimulation reward" is marred by a meagre half-page of theoretical discussion of incentive motivation, coming almost as an afterthought. Sadly, this approach encourages the present tendency of "behavioural neuroscience" to be a subject driven by phenomena rather than by thought.

A third temptation is to litter the text with trendy material, perhaps not only to capture the excitement of contemporary research but also to make the opposition obsolete. Brown and Wallace have rightly introduced much interesting recent material, for example, on enkephalins and pain, the neurochemical substrates of learning, and, impressively, on "brain grafts". However, the "shock of the new" is not always right. What is exciting today may be reinterpreted or not even replicated, tomorrow. Although it is perhaps healthy for students to hear of controversy, it is dangerous to present it as fact. In addition, driven by the thirst for novelty, the reader expects updating in every topic, whereas Brown and Wallace's account of dopamine levels in schizophrenic brain neglects important work.

These are perhaps harsh criticisms of what is a well-produced and solid textbook. The choice and clarity of the figures is particularly good, and the writing emulates in readability that of Carlson's *Physiology of Behavior* (Allyn & Bacon: 1977), although in my opinion the latter is a more successful textbook.

In contrast to *Physiological Psychology*, Cotman and McGaugh's work represents a significant change of style in textbooks on the subject. Although much of the material overlaps with the former book, the title *Behavioral Neuroscience* emphasizes its preoccupation with neurobiology, which I found to be the strongest feature of the book.

The authors, while compiling a vast fund of useful information, have attempted to solve the problem of its critical assessment by enlisting about 20 specialist collaborators. Unfortunately, despite this expertise, the book fails to achieve that essential thematic coherence and balance in level of presentation. Some of the chapters, such as those on learning and memory are very detailed, whereas others, for example on thirst and hunger, are not. The superior colliculus is barely mentioned in the chapters "Seeing" and "Moving", and neuropsychological topics such as functions of association cortex and frontal lobe, and aphasia, are given short shrift.

The book is excellently produced with one glaring exception; the many sketches accompanying the text are poor substitutes for scientific drawings, and in a few cases, are embarrassing and superfluous.

Can there then be an ideal textbook in physiological psychology? In preparing my own courses I have found none, although many of the textbooks are very useful. If we eschew the possible unfortunate development of textbooks being written (and reviewed by) committees, then a better approach would seem to build a modular course by using books on specialist topics, such as R. Melzack's excellent and inexpensive *The Puzzle of Pain* (Penguin, 1973).

Superficial inspection of Kolb and Whishaw's *Human Neuropsychology* suggests that it will be a valuable weapon for the latter armoury, especially since it covers areas only poorly represented in the other books reviewed. I was impressed with the attempt to integrate both theory and human and animal experiments in sufficient depth to provide adequate coverage of subjects such as the neural mechanisms of language.

Although this book will be invaluable for advanced courses in physiological psychology and clinical psychology, I doubt it will be the last word on textbooks in neuropsychology. However, I also doubt that we have seen the last textbook of physiological psychology. ☐

Trevor W. Robbins is a Lecturer in Experimental Psychology at the University of Cambridge.

More sight than sound

Stuart Sutherland

Perception: The World Transformed. By L. Kaufman. Pp.416. (Oxford University Press: 1979.) Hbk £10.75; pbk £6.95, $14.95. *Perception: An Applied Approach.* By W. Schiff. Pp.493. (Houghton Mifflin: 1980) $18.50, £9.95. *The Psychology of Visual Perception.* 2nd Edn. By R.N. Haber and M. Hershenson. Pp.431. (Holt, Rinehart & Winston/Holt Saunders: 1980.) $22.95, £10.50.

TWENTY years ago there was no worthwhile textbook on visual perception. This deficit was removed by the publication in 1975 of Kaufman's *Sight and Mind* and Rock's *An Introduction to Perception*, the former an advanced-level undergraduate textbook, the latter rather more elementary, but both excellent. We now have three further textbooks on vision — despite their titles, *Perception: The World Transformed* and *Perception: An Applied Approach* deal almost entirely with vision and both treat audition in a cursory and inadequate fashion.

Of the three new books, Kaufman's is the easiest to read, but is less up to date and is considerably more elementary than the other two. Haber and Hershenson's *The Psychology of Visual Perception* is exceptionally lucid and gives fair and unbiased accounts of the theories with which it deals. Schiff's text is both the least conventional and the most scholarly: as well as summarizing our current knowledge of perception, it shows how this knowledge can be applied to such problems as the design of aids for the blind and the legibility of text, and its extensive references provide the reader with ready access to the recent literature on vision.

The books have different virtues, but they all share the same fault. The mechanisms of visual perception are the way they are because they have evolved to carry out a particular task, namely, to reconstruct from the fragmentary two-dimensional data in the retinal image a three-dimensional model of the environment and to recognize the objects represented in that model. Although all the authors acknowledge that vision has a function, they fail to relate the details of the research they describe to that function. Thus, they describe the Gestalt psychologists' laws of Prägnanz (for example, the tendency to group together similar visual elements arranged in a straight line or smooth curve) without giving any indication of how the application of these laws enables the visual system to carry out its primary function.

Moreover, none of the books gives any account of the computer models that have been devised over the past decade to extract three-dimensional representations from a visual image. The attempt to construct such models has revealed the nature of the

Bring Your
Library Up-To-Date
With This ANS Edition...

Paul Cohen
MONOGRAPH

Water Coolant
Technology of
Power Reactors

456 pages — 6″ × 9″
Hardbound $27.90

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for ____ book(s).
☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name ____
Organization ____
Street ____
City ____
State ____ **Zip** ____

Circle No. 15 on Reader Enquiry Card.

problems that the visual system must solve, and in many instances has suggested how it actually does solve them. Not one of the books makes any reference to the work of the late David Marr, perhaps the only genius to work on the psychology of vision since Helmholtz himself. Marr's work in fact started where Helmholtz stopped: Helmholtz rightly insisted that visual perception was mediated by a series of complex inferential processes, but he was unable to specify with any rigour what these processes were. The existence of the digital computer has allowed Marr and his associates to construct computational models of the visual system that have thrown a completely new light on how it must function. Unfortunately, his work is not easy to follow and too few psychologists have grappled with it: the only good popular account is to be found in J.P. Frisby's *Seeing: Illusion, Brain and Mind*.

Two of the publishers of the three books under review have served their authors badly. In *Perception: The World Transformed* some of the figures are unnecessarily hard to interpret (e.g. p.128), others have incorrect legends (e.g. p.190), and at least one name that appears in the index does not appear on the specified page or indeed anywhere else in the text. The sins of Oxford University Press are, however, comparatively venial compared with those of Houghton Mifflin. *Perception: An Applied Approach* is the worst designed book I have ever encountered. It can be read only with extreme difficulty since each page contains three columns of text with no justification of the right-hand margins; much of the information purveyed is contained in those horrid boxes so popular with American publishers, but Houghton Mifflin have plumbed new depths of idiocy by spreading a single box over as many as five different pages, with areas of normal text and figures spattered randomly around its edges; as though these faults did not make the reader's task sufficiently difficult, the book's designer has made it almost impossible to associate the figure legends with the relevant figure, even going to the length of placing figure and legend on different and non-facing pages. The format of the book is a triumph of bad design, a nice irony given that the book's content is intended to show how a knowledge of vision can improve the design of artefacts.

In summary, then, *Perception: The World Transformed* is an excellent introduction for the novice; *The Psychology of Visual Perception* is a good but conventional treatment and is nicely complemented by *Perception: An Applied Approach*, which includes much material not normally found in undergraduate textbooks. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Information in mind

Angus Gellatly

Cognitive Psychology. By R.L. Solso. Pp.499. (Harcourt Brace Jovanovich: 1980.) £10.95, $17.95. *Cognitive Psychology and Its Implications*. By J.R. Anderson. Pp.503. (W. H. Freeman: 1980.) £8.10, $15.

IN THE preface to his book, Robert Solso identifies two key questions raised by the information-processing approach to cognition: what are the stages through which information is processed? and in what form is information represented in the human mind? Although each of these texts addresses both questions, they differ in the emphasis given to them.

Anderson's major concern is the representation of knowledge, and much of his book is an exposition of recent models of long-term memory that employ a propositional format; indeed, he sees such models as the major achievement of modern cognitive psychology. The organization of the book reflects his commitment; peripheral stages of perception and attention are covered in a single 40-page chapter, after which interest turns to the evidence on knowledge representation. Only when his conclusions on this issue have been established does Anderson move on to apply them to the perennial problems of learning, memory, reasoning and language. By contrast, Solso's commitment is more to the methods of information-processing psychology than to a particular theoretical stance within that framework. His is the more traditional approach, starting with the detection of sensory signals and charting the flow of information into the processing system. Over 100 pages are devoted to reception and early analysis of stimuli, with consideration of memory in its varied manifestations forming the second part of the book, and higher-order cognitions the third. Both authors display an admirable familiarity with the vast literature. Anderson's clearly defined viewpoint lends his book a vigorous proselytizing air and a unity of purpose, but Solso, by allowing greater space to studies published earlier than the past decade, suggests a longer-term perspective. Anderson provides particularly good chapters on mental imagery, schemas, cognitive skills and deductive reasoning. Solso is also good on mental imagery and on attention, and he scores over Anderson by including a chapter on cognitive development.

On the debit side, the two books are open to a number of the same criticisms. Both are concerned almost exclusively with what cognitive psychologists do in the laboratory using undergraduate subjects and computers, rather than with cognition in any broader sense. Neither author makes more than passing reference to relevant neuropsychological findings, to the issues of individual and cultural differences, or to applied studies of reading. Perhaps, granted the assumptions they make, these are allowable omissions, but both authors advance staggering claims for future developments and applications of cognitive psychology that are in stark contrast to the well-worn examples of such applications that they can cite.

For teaching purposes, what are the relative merits of the two books? Solso perhaps provides the more general and easier introduction to cognitive psychology, Anderson is for the second- or third-level option course. □

Angus Gellatly is a Lecturer in Psychology at the University of Keele.

Entomological ethology

Stewart Evans

Introduction to Insect Behavior. By M. D. Atkins. Pp.237. (Macmillan, New York/Collier Macmillan, London: 1980.) Pbk $9.95, £6.50.

ALTHOUGH the value of studies in invertebrate behaviour is widely accepted, these animals still tend to be neglected in textbooks of ethology. Some general texts, such as those by Alcock (*Animal Behavior: An Evolutionary Approach;* Sinauer, 1975) and Manning (*An Introduction to Animal Behaviour;* Edward Arnold, 3rd Edn, 1979), do strike a satisfactory balance between different animal groups but others give the impression that almost all of the important research is on birds and mammals.

Clearly, students need access to a wider coverage of the behaviour of invertebrates, and Atkins's book goes a long way towards providing this for insects. In an admirably clear account, he concentrates on functional aspects of behaviour and, with well-chosen examples, illustrates both its complexity and variability. I believe that undergraduates will enjoy reading the book; they will find it informative, up-to-date and, in parts, dynamic. For instance, the chapter which includes a consideration of genetic aspects of behaviour underlines the way in which both traditional ethological techniques, such as those employed in Rothenbuhler's studies of hygienic and unhygienic honey bees, and entirely new ones, such as those involving the use of *Drosophila* mosaics, are making inroads in this area of research.

One major disappointment is, however, that the author almost totally ignores the work of insect neurophysiologists. It is difficult to understand how Huber's research, or that of several others in the same field, could be overlooked. Students ought to be aware of this work and, in my view, the book suffers from missing the opportunity to include it. This omission apart, I think that the book is good value for money. It is attractively illustrated and provides a good reading list; hopefully students will be stimulated to make extensive use of it. ☐

Stewart Evans is a Senior Lecturer in the Department of Zoology at the University of Newcastle upon Tyne.

Diversity of insects

R. W. Thorp

Entomology. By Cedric Gillott. Pp. 729. (Plenum: 1980.) Hbk £31.19, $49.50; pbk £14.18, $22.50.

THIS book is a comprehensive introductory college text for a senior undergraduate class in general entomology. It is designed to contrast with the "traditional taxonomic approach" which provides little information on insect physiology and ecology.

The 24 chapters are grouped into four major divisions: evolution and diversity; anatomy and physiology; reproduction and development; and ecology. The first of these comprises about 45% of the book, and deals with insect diversity, arthropod evolution, and classification and identification of insects, and includes a chapter on external structure. A key for the identification of all orders is provided; description of each order is followed by discussions of structure, life history and habits, and phylogeny and classification, including brief mention of the more common or important taxa. Most of the rest of the book, divisions 2 and 3, deals with structure and function of the integument, and all internal systems, and with embryonic and post-embryonic development. Finally, the author considers how insects relate to their abiotic and biotic environments, and concludes with a chapter on applied entomology entitled "Insects and Man".

The strength of the book is in the areas of physiology , developmental biology and the ecological relationships of these to abiotic environmental factors. Also, the chapters on arthropod evolution, insect diversity, classification, and identification at and above the ordinal level, especially the phylogenetic discussions, are more complete than in most single-authored texts.

Weaknesses include the lack of a discrete section on insect behaviour, and inadequate discussions of population dynamics and life tables, forest entomology, zoogeography and community ecology, thermoregulation in insects, interactive effects of abiotic factors, host specificity, economic thresholds, and methods for insect population management. The section on chemical control is poor in that possible resistance to insect growth regulators is not discussed, nor are the potential uses of anti-juvenile hormone compounds, antifeeding compounds and chemosterilants. The absence of keys and descriptions of infra-ordinal taxa reflects the author's feeling that identification is best relegated to laboratory exercises and to collections made by the students; however, this means that an additional text with keys to and descriptions of families will be needed for the lab.

It may be concluded from the author's criticism of introductory texts which are taxonomically oriented, and from my criticisms of the gaps in his coverage, that it is not easy to find a suitable text to accompany both lectures and practical work in entomology. Although most authors agree that a balanced coverage is needed, one is quickly overwhelmed by the vast amount of material to be covered in a quarter or semester. After the fundamentals of the subject have been covered, there is often little time left to introduce the fascinating diversity of natural history data available. Thus, the "balance" achieved in the introductory texts and courses currently available seems to differ according to the bias of the author or instructor. No two texts ever strike the same balance, but perhaps this is not as important as the development of student interest in this fascinating field. This new text by Dr Gillott succeeds in accomplishing this and is one of the better books available for an introductory entomology course. ☐

R. W. Thorp is Professor of Entomology at the University of California, Davis.

A new "Life" for vertebrates

Barry Cox

Vertebrate Life. By W.N. McFarland, F.H. Pough, T.J. Cade and J.B. Heiser. Pp.875. (Macmillan, New York/Collier Macmillan, London: 1979.) Hbk $29.95, £12.95; pbk £6.95.

THIS book fills a very great gap in the books that one can recommend to students. There are several adequate, sound books on the comparative morphology of vertebrates, and reference to them is certainly necessary if the student is to understand the details of the evidence of their structural and adaptive radiation. Nevertheless, with their lengthy exposition of how the hip bone is connected to the thigh bone, or of the position of the metaconule, such books provide little stimulation to the enquiring mind. Instead, they encourage the poorer student's tendency to cram facts as an alternative to the more demanding activity of trying to understand principles and to solve problems.

As the authors explain in their preface, this book is built around two themes; the evolution and the ecology of vertebrates, dealt with in a way that also integrates relevant aspects of their physiology, behaviour, locomotion or morphology. So although the list of chapter titles includes the familiar groups in the normal phylogenetic order, many other topics are either covered in separate sections within the chapters or make up independent chapters.

The book is divided into 22 chapters, half of which are by a single author (Cade or Pough), while 9 were written by McFarland and Heiser. The chapters seem uniformly to be really up to date in their approach and in the references given at the end of each. An average of only eight references per chapter is rather low, though some others are given after the illustration captions. The illustrations themselves are excellent and well chosen. Many of them have been redrawn, to provide a pleasing similarity of style thoughout the book. In some cases the authors have enlarged on a particular topic (for instance the evolution of amphibian vertebral structure) by means of a lengthy commentary-caption, so that the flow of the main text is uninterrupted.

After a brief introduction to the basic body plan of the vertebrates, the second chapter discusses their ancestry and environment of origin. The third includes very useful sections on the theory of classification and its application to vertebrates, and on species and speciation. The chapters on the different vertebrate groups deal with extinct as well as with the living groups, and with such additional topics as the origin of land vertebrates, adaptations to life in the deep sea, the ecology and physiology of dinosaurs, and the waves of extinction at the end of the Cretaceous and during the Pleistocene Ice Ages. The authors are careful to emphasize that each group possesses abilities and adaptations in its own right and that, for example, ectothermy has positive value in some ways of life rather than merely being a characteristic of groups that "failed" to become endothermal.

Five chapters scattered through the book give useful summaries of the positions of the continents and of the climatic regimes at different periods of the Earth's history. Other chapters deal with the influence that life in the water has had on basic vertebrate functions, and on the adaptations of endotherms to rigorous habitats.

In addition to being well-written and well illustrated, the paperback edition is extremely good value. It should put new life into any vertebrate course, and it thoroughly deserves to capture the bulk of the market for vertebrate textbooks. □

Barry Cox is Professor of Zoology at King's College, University of London.

Fossils the French way

John C. W. Cope

Elements of Palaeontology. By Claude Babin. Translated by N. Orriss and edited by D. Palmer. Pp.446. (Wiley: 1980.) Hbk £19, $57; pbk £7.95, $22.50.

IN HIS introduction to this book the editor mentions the important point, certainly not implicit in the title, that there is an "assumption that the reader is not a novice to palaeontological and biological terminology and technicalities". The approach therefore differs considerably from Rhona Black's *The Elements of Palaeontology* (Cambridge University Press, 1970), a book which has rapidly become the standard introductory palaeontological text.

To the British student the first part of the book will prove to be the most useful. The first chapters cover fossilization, fossil collection and preparation and the basics of palaeontological taxonomy. These are adequate treatments of the subject, but the exhortation to "gather the greatest possible number of samples" will not be welcomed by those with an eye on conservation. There follows a useful résumé of palaeoecological methods; this is well illustrated and many examples will be familiar to English-speaking students.

Part Two of the book carries the curious title "Some Facts of Evolution". After consideration of the origin of life, various aspects of fossil plants and some invertebrate fossil groups are considered. Although the treatment is claimed to be selective and not morphological or taxonomic, I find it difficult to understand why major groups such as sponges, bryozoa or gastropods are omitted whilst minor groups such as hyolithids, conulariids and tentaculitids are discussed. Treatment of major groups is often idiosyncratic. Under "Echinoderms", for instance, a short discussion of pentaradiate symmetry is followed by a presentation of the stylophoran/calcichordate controversy. There is no mention of the other 15 or so echinoderm classes — several of which could provide first-rate examples of evolutionary trends. Similarly, treatment of bivalves is restricted to their origins and their Ordovician diversification. The same approach is applied to other groups, an approach which, rather than stimulating, frustrates by its inadequacies. The final chapter in this section contains some 60 pages on aspects of vertebrate evolution, including a substantial section on hominid evolution.

The final part of the book, "Palaeontology and Evolution", briefly discusses some ideas on evolutionary theory.

The translation (from French) is acceptable but rarely more; it contains a few glaring errors (for instance that the bivalve ligament closes — instead of opens — the shell) but peculiar phraseology and word-order frequently make it awkward to read. Even with an English dictionary to hand, a few sentences are totally incomprehensible.

The good illustrations and useful reference lists cannot compensate for the sketchy and disjointed coverage in Part Two of the book, and although it contains much that is interesting, I must say that I found this book a disappointment. □

John C. W. Cope is Senior Lecturer in Geology at University College, Swansea.

Micropalaeontological survey

Alan Lord

Microfossils. By M. D. Brasier. Pp.193. (George, Allen & Unwin: 1980.) Hbk £12, $27.50; pbk £6.50, $14.95.

THE growth and development of micropalaeontology has been closely connected with interest in the world's oceans, coupled with its application to hydrocarbon exploration. The results of the *Challenger* Expedition (1873–1876) stimulated the study of small organisms and their role in modern sedimentation. Stratigraphical studies of fossil counterparts by such pioneers as T. R. Jones in Britain eventually led to the recognition in the 1920s of the use of microfossils for correlation, particularly between boreholes where drilling usually destroyed macrofossils. Since that time various groups of microfossil have been found to be useful for biostratigraphical and palaeoenvironmental purposes, usually with industry recognizing the value of particular fossils in particular situations after initial academic study. The continuing exploration of the ocean floors by D/V *Glomar Challenger* has depended on microfossils for stratigraphical control.

Dr Brasier's book provides a broad survey of microfossil groups, giving information for each on biology, hard- and soft-part morphology, ecology and distribution, relationship to sedimentation, classification and applications, together with suggestions as to further reading and hints for study. Coverage of each group tends to reflect its past importance to geologists — foraminifera, ostracods, spores and pollen, conodonts, dinoflagellates and coccoliths are given space in that order. Also included are a brief introduction, a useful appendix on techniques and a good basic bibliography. The style is economic, without a great burden of cited literature, but with many pointers to the student for further information and enough named examples to round out each chapter. Numerous line drawings illustrate the book; use of these is clearly a matter of personal preference when excellent photographs of microfossils are available nowadays.

The book is thoroughly recommended for undergraduates and junior postgraduates, not least because it complements the more detailed *Introduction to Marine Micropaleontology,* edited by B. U. Haq and A. Boersma (Elsevier, 1978). The paperback is certainly a good buy for anyone with an interest in or need to know about microfossils. □

Alan Lord is a Lecturer in Geology at University College London and teaches part of the MSc course given in the Postgraduate Unit of Micropalaeontology.

Riches of biology

E. C. Cox and R. M. May

Biological Science. 3rd Edn. By W. T. Keeton. Pp.1080 + glossary and index, pp.A51. (W. W. Norton: 1980.) $19.95, £11.50. Study guide $4.95, £2.50; teacher's manual no charge.

TEACHING introductory biology is not easy. An average class has an extraordinarily diverse set of interests, preparations, aptitudes and appetites for the subject. It is never clear, moreover, where one should start and where, or at what level, one should end, given both the richness of the subject matter and the diversity of the audience. These difficulties surface in the pages of many introductory texts, often reflecting all manner of confusions on the part of the author about pace, depth, coverage and readership. The problems are variously solved by dealing only with animals or plants; by reducing everything to a few weak bonds, permuted; by centring the discussion on the fate of mankind; or by omitting, except *en* *passant,* most of the evolution of structure and function by natural selection. Each of these narrow strategies, when vigorously pursued, can appeal to a particular audience. There is no doubt, for example, that a student already well prepared in the natural sciences can, upon arrival at a university, understand and benefit from Watson's *Molecular Biology of the Gene* or Lehninger's *Biochemistry,* two texts normally used by upperclassmen. But there are few such students. Most of us need, and some even want, a broad and responsible introduction to biology taught in a serious way at a high level. Fortunately for us, Keeton's third edition is now available.

There are many features of the new edition that set it apart from its competitors and represent improvements over the second edition. Keeton's approach has always been to begin with the rudiments of chemistry, proceed to the structure and function of macromolecules, cell structure and function, and thence on through behaviour, genetics and evolution, community ecology, the origins of life and the description of the phyla. He covers each topic thoroughly, judging the depth and breadth with care, and adding the proper notes of scholarly caution and scepticism. The book has an air of authority; too many introductory texts (and lecturers) substitute glib metaphor for real explanation (see our earlier reviews in *Nature* **284,** 106; 1980). The text consequently does not always make for a good read, but since it is aimed mainly at students in the natural sciences, and at those graduate students (and their teachers) reviewing areas outside their special competence, this cannot be faulted. No one could be expected to remember everything in this book; indeed, Keeton says in the introduction that he obviously could not remember it all himself.

The broad theme running through Keeton is that evolutionary theory accounts for adaptive strategies. It is the thread that ties together the material on the cell, energy transformation, the biology of whole organisms and, of course, genetics, evolution, behaviour and population biology. This not only gives students some feeling for the deep structure of the discipline but, as important for a beginner, supplies the outline of an image whose shape is finally supplied by what at first sight is an incoherent body of detail, often alarming in its variety.

The prose is enhanced enormously by first-rate illustrations and photographs, full colour as well as black-and-white. Many examples could be given. Two that stick in our minds are the drawings combined with electron micrographs to illustrate typical plant and animal cells with pleasing clarity, and the beautiful colour photographs of the world's major biomes and of plant pollinators in the act.

Inserts on special topics are also used; this is now a fairly standard pedagogical ploy, but it is nonetheless effective, for it allows small essays on advanced or special topics that would otherwise interrupt the text.

A list of test questions and a teacher's manual accompany the text. The former, in our view, is useful but not exceptional. The latter we can recommend, especially to lecturers who have not previously taken on an introductory course. It has several strengths, not the least being excellent advice on how to pose examination problems and on which of the areas covered in the lectures most often confuse the students.

Keeton's untimely death last year was a sad loss. This excellent text is likely to stand as a memorial for some time. ☐

Edward C. Cox is Chairman of, and Robert M. May Professor in, the Biology Department at Princeton University, where they jointly teach an introductory biology course.

Maths applied

P. T. Saunders

Statistics and Experimental Design. 2nd Edn. By G. M. Clarke. Pp.188. (Edward Arnold/University Park Press: 1980.) Flexi £6.50, $24.50. *Mathematics and Statistics for the Bio-Sciences.* By G. Easen, C. W. Coles, and G. Gettinby. Pp.578. (Ellis Horwood: 1980.) Hbk £25, $90; pbk £8.50. *Statistics for Biologists.* By D. J. Finney. Pp.165. (Chapman and Hall/Methuen: 1980.) £3.75, $9.95.

WHEN the first edition of *Statistics and Experimental Design* appeared, *Nature's* reviewer wrote that it was "clear and well written and may be strongly recommended to non-mathematical students requiring a common-sense grounding in statistical methods". I entirely agree with this judgement. The explanations are lucid, and well illustrated with examples. There are many exercises, for which model solutions are provided. The author has also added a liberal sprinkling of comments which contribute towards an understanding of the methods he is describing and which should help the reader to avoid some of the pitfalls that await the unwary user of statistics.

The contents are not much changed from the first edition. They include an introduction to probability theory, tests of significance (both parametric and non-parametric), correlation and regression, the principles of experimental design, and an account of factorial experiments including corrections for missing observations and non-normality. Students and research workers who can master the material in this not-too-large book should be well equipped to begin using statistics; those who cannot might be better advised to leave statistics alone.

Statistics for Biologists covers a similar range of material, but deals with fewer topics. This is partly because it is shorter, and partly because the author is more concerned with introducing the principles of statistics than with providing instruction in how it is used.

To some extent he succeeds, but I feel he often devotes to comment and discussion space which could more profitably have been used for techniques, even with his stated aim in mind. For example, regression occupies over four pages, yet the author neither explains the principle of least squares nor gives the formula for fitting even a straight line. Surely this must reduce the chance that the student will understand regression, if only because he cannot be given numerical examples to work out.

As an indication of the level to which students are taken, the author writes that the distinction between one- and two-tail tests is "logically rather subtle" and suggests that the reader should ask a statistician to explain it to him later on. I would have thought that this was one of the basic principles that students ought to learn; it is even hard to read many statistical tables if you don't understand this point. The problem does not arise within the book, as no full tables are provided. Students who find the author's style to their taste will profit from reading this book, but it will leave them with a lot more to learn.

If *Mathematics and Statistics for the Bio-Sciences* had been better written, it could have helped to satisfy a real need. But there are too many things wrong with it. The authors have chosen to write in a rather formal style, which they do not always handle well. I was not impressed with their explanations of such important ideas as limits and derivatives; these are difficult concepts for beginners to grasp, and require a better and more carefully worded treatment than they receive here.

The presentation is often slipshod. For example, while it is certainly true that "Two matrices **A**, **B** can only be equal to each other if they are the same size, say $m \times n$, and all their corresponding elements are equal", the authors should have made it clear that this is the *definition* of equality, and that there is not, as their statement implies, some further condition to be satisfied. It also seems odd to introduce the Poisson distribution and even calculate its mean and variance without mentioning the assumptions on which it is based and (therefore) its main application.

The one really good feature of the book is the large number of realistic biological examples and exercises. Anyone teaching mathematics to biologists will find these very useful. But I cannot recommend the book to students, either for classroom work or for individual study. ☐

P. T. Saunders is a Lecturer in Mathematics at Queen Elizabeth College, University of London.

Across the span of biophysical chemistry

Joseph E. Coleman

Biophysical Chemistry, in 3 parts. By C. R. Cantor and P. R. Schimmel. Pp.1,370. (W. H. Freeman: 1980.) Pt I hbk £17.70, $31.95; pbk £9.20, $16.95. Pt II hbk £20.70, $37.95; pbk £10.60, $19.95. Pt III hbk £22.40, $39.95; pbk £11.90, $21.95.

IT IS difficult to give a simple definition of biophysical chemistry. Scientists who have considered themselves biophysicists have covered such a wide variety of subjects, from the physics of ion transport to crystallography, that it is difficult to settle on a satisfactory definition. To avoid narrowness, it is perhaps best to define the field as that in which physical methods are applied to the study of biological systems — molecules as well as larger aggregates. To date, such an immense number of physical techniques have been used in this way that it has been difficult to find a textbook which covers even a representative number of them. At the same time, the study of these methods must be coupled to a reasonably complete understanding of the structure and function of the objects under study — in most instances biological macromolecules.

Cantor and Schimmel have chosen to attack the problem of the breadth of biophysical chemistry by producing a three-volume work of nearly 1,400 pages. The concept is admirable. The first volume, entitled *The Conformation of Biological Macromolecules,* sets the stage by discussing in detail the structures to be studied by various physical methods. Here, the authors have in fact set one of the few limits to their scope by confining the discussion for the most part to homogeneous macromolecules. While the existence of macroassemblies is mentioned in appropriate places, this limitation seems reasonable in that most of the best illustrations of the applications of biophysical chemistry come from studies of homogeneous biological macromolecules. The authors on occasion indicate the power of the methods of studying subcellular particles or even whole cells, but this aspect is not emphasized. Hence the first volume is devoted primarily to discussion of protein and nucleic acid structure, polysaccharides, lipids and membranes being given rather less space. The treatment is thorough, starting with composition and chemistry of the constituent monomers and extending to the static structure of the polymers and on to a discussion of the forces that determine protein folding and nucleic acid structure.

Having set the stage by presenting the structures to be studied, the second volume, *Techniques for the Study of Biological Structure and Function,* presents the physical techniques required to determine the macromolecular structures. Since the methods covered are those used to derive the information in the first volume, a purist might quibble about the relationship of the cart to the horse, but as a learning format such an approach has an advantage, since the context for the application of these (bio)physical methods is clear beforehand. The number of physical techniques applied to one or another biological macromolecule has become so great that it is difficult to be all-inclusive, and Vol. II is not. In a number of specific areas the presentation is extensive, as for all types of optical spectroscopy; hydrodynamic techniques are also well covered. Approximately 130 pages are devoted to X-ray crystallography, including scattering, and to a short mention of extended X-ray absorption edge spectroscopy. This is a thorough and well-presented discussion.

In the magnetic resonance section, both NMR and EPR are covered. The EPR discussion is rather brief; no examples of transition metal EPR spectra are given and the student will have difficulty in understanding even the most basic aspects of the interpretation of EPR spectra from the material given. On the other hand, the NMR section gives an excellent presentation of the theoretical and basic aspects, including relaxation mechanisms. The applications section is not as good. Most of the spectra are out of date; NMR technology has improved so much that even ^{1}H spectra from the early 1970s do not do justice to the field.

The third volume is entitled *The Behavior of Biological Macromolecules.* The idea behind this section would appear to be that since the student now knows his structure and the methods for discovering it, he can now use this information to understand some of the more complex interactions of biological molecules. These include ligand binding to macromolecules (both equilibrium and kinetic aspects), the variety of "allosteric" phenomena involved in biological regulation, and conformational equilibria found in proteins, nucleic acids and membranes. Many specific examples of the various methods dealt with in Vol. II are given in Vol. III and make the discussion of applications of biophysical techniques more complete. This last volume also covers some other methods such as differential scanning calorimetry — an important addition. The extensive ^{1}H-NMR studies of t-RNA are presented here with spectra somewhat more up to date than those appearing in the NMR section proper.

The book is sufficiently mathematical to be relatively rigorous without overdoing it. There is an appendix giving a brief review of matrix algebra, and appropriate problems are found at the end of many of the chapters, with answers in the back of each book.

Minor criticisms aside, one has to conclude that the three-volume concept of presentation was well-founded and that the execution is excellent. I believe that this will be the textbook of choice for a course in biophysical chemistry; I can think of no other text that is as complete or as consistently well done over such a broad spectrum of subjects. □

Joseph E. Coleman is Professor of Molecular Biophysics and Biochemistry at Yale University.

Biochemical quartet

David Saggerson

Biochemistry: A Case Oriented Approach. 3rd Edn. By R. Montgomery *et al.* Pp.711. (C.V. Mosby: 1980.) Pbk £11.50, $19.95. *Biological Chemistry: The Molecular Approach to Biological Systems.* By K.E. Suckling and C.J. Suckling. Pp.381. (Cambridge University Press: 1980.) Hbk £25, $59.50; pbk £9.95, $19.95. *Biochemistry and Physiology of the Cell: An Introductory Text.* 2nd Edn. By N.A. Edwards and K.A. Hassall. Pp.448. (McGraw-Hill: 1980.) Hbk £15.95, $38.20; pbk £7.95, $19. *A Guidebook to Biochemistry.* 4th Edn. By M. Yudkin and R. Offord. Pp.261. (Cambridge University Press: 1980.) Hbk £18, $53.50; pbk £6.95, $15.95.

OVER the past 10 to 15 years, many excellent biochemistry textbooks have appeared. To succeed in the face of such competition, any new book must be outstanding or have some unique and unusual features.

The general textbook by Montgomery *et al.* fits both of these criteria. As its title and the sketches of clinicians at work on its cover imply, it is aimed to serve the needs of students of medicine or other health-related subjects. Each chapter first presents a particular set of topics in much the same way as other, more conventional textbooks of biochemistry. This is then followed by discussion of various clinical conditions, questions and problems. The number and range of conditions covered in the book is impressive. The order of chapters is rather unconventional, starting with nutrition, followed by proteins and enzymes and the acid–base balance, before discussion of cellular metabolism and its regulation. It is explained that the introduction of these particular topics early in the book permits a wider range and fuller discussion of clinical cases in subsequent chapters. I cannot argue with this. The book is imaginative, full of information, well referenced (up to 1979 in some cases) and clearly set out. As someone who is personally involved with the challenge of imparting biochemical knowledge to medical students in an interesting and relevant fashion, books of this kind are most welcome. It is likely, though, that the book requires a course to be deliberately constructed around it (as the authors do at the University of Iowa College of Medicine).

The Sucklings' book falls clearly in to two parts. The first two-thirds is largely devoted to explaining in chemical mechanistic terms a wide range of biochemical reactions. Since there are few books with this approach, this is useful either to explain biochemistry to chemists or chemistry to biochemists. Each chapter contains a number of problems which are solved with brief explanation at the end of the book; however, it was not always apparent to me that the information presented was sufficient to enable a student to tackle the problems satisfactorily. The last third of the book covers a number of topics such as enzymes as catalysts, enzyme mechanisms (using kinase-catalysed reactions as examples), control of enzyme activity, biochemical techniques (separation methods and structure determination), membrane structure, drugs (design and metabolism) and communication between cells (hormones and the immune response). It is not clear why these topics have been selected for inclusion considering the nature of the earlier part of the book. This curious mixture is presented in a rather superficial way. One cannot but feel that the book would have been better if this last third had been omitted and the topics left for the general textbooks where they are better covered.

Cellular Biochemistry and Physiology by Edwards and Hassall (McGraw-Hill, 1971) has now metamorphosed into *Biochemistry and Physiology of the Cell*. This is similar in size to its predecessor, having the same objectives presented in the same style. In a modest-sized book the authors have managed to incorporate a considerable amount of material and present it in a clear and concise manner. I see this book as being a valuable introductory text for biochemistry students or as a main text for those following biochemistry as a subsidiary subject. The main areas of the subject are dealt with — protein chemistry, enzymes and their kinetics, intermediary metabolism and protein synthesis, together with specialist sections on nerve, vision and muscle contraction — and throughout the book a number of important techniques are introduced and described. I found the discussion of the control of metabolism, including hormones, which has been considerably expanded since the 1971 edition, to be particularly good. As the authors themselves say, increases in knowledge rendering a book "more advanced" may in fact make it easier to read. Incidentally, this book was the only one of the four reviewed here presenting a credible, but simple explanation of the chemiosmotic theory of oxidative phosphorylation.

The book by Yudkin and Offord is the most elementary of the four. In general it presents a superficial outline of biochemistry under the four section headings, "Structure and Function of Macromolecules", "Metabolism", "Molecular Genetics and Protein Synthesis", and "Compartmentation and Regulation". With 250 rather small pages it cannot cover the same amount of ground as Edwards and Hassall; in fact some important topics are missed out and the explanation of others is rather thin. The book also suffers from a rather dull format. I was particularly disappointed to find no discussion of any of the important techniques of biochemistry nor any simple enzyme kinetics; this latter omission makes explanation of some aspects of enzyme action difficult. I also feel that the authors' approach to intermediary metabolism would have benefited from the use of more diagrams and from an attempt to present it within a physiological framework. Metabolism is much more interesting if metabolic specialization of different tissues is explained alongside the physiological conditions which favour particular metabolic processes. ☐

David Saggerson is Lecturer and Tutor to Medical Students in Biochemistry at University College, University of London.

Molecular cytology

John Edwards

Cell and Molecular Biology. 7th Edn. By E. D. P. De Robertis and E. M. F. De Robertis, Jr. (Holt, Rinehart & Winston/Holt Saunders: 1980.) $24.95, £10.95.

Is cell biology a discipline in its own right? What, for example, are the prospects for a biological degree course which cuts horizontally across the traditional divisions between animals, plants and microbes, as does biochemistry, but focussed at the level of the cell? Textbooks with "cell biology" in their titles often are heavily weighted towards biochemistry, usually shorn of its tougher shadings into physical and organic chemistry. Not so with this volume, evolved over more than 30 years from a text of general cytology.

The core of this substantially rewritten edition continues to treat the cytoplasmic organelles and nuclei of eukaryotic cells, bringing together ultrastructural, cell fractionation and functional investigations, and is well illustrated and usefully documented (more than a third of the references post-date the previous edition). Mainstream molecular biology (one-sixth of the whole) leads via cell differentiation

Advanced Book Program

MOLECULAR ELECTRONIC STRUCTURES
An Introduction*

Carl J. Ballhausen
University of Copenhagen

Harry B. Gray
California Institute of Technology

An introductory account of the electronic structures of atoms and molecules with emphasis on the nature of the chemical bonding in various types of molecules. After completing this text the physical chemistry student will be able to handle polyatomic inorganic and organic molecules as well as the usual diatomics and triatomics. The book is useful also for advanced inorganic chemistry courses and as a self-study unit for the chemist who needs to keep current with this rapidly moving area of modern chemistry.

Contents:

THE SCHRÖDINGER EQUATION
Formalism, The Particle in a Box, Atomic Orbitals, The Electron Spin

MOLECULAR ORBITALS
Diatomic Molecules, Symmetry Operators

MANY-ELECTRON SYSTEMS
The Pauli Principle, Atomic Electronic Configurations, and Energies

ELECTRONIC STATES OF MOLECULES
Hybrid Orbitals, Homonuclear and Heteronuclear, Diatomic Molecules, Water, Carbon Dioxide, Formaldehyde, Ethylene, Allene, Diborane, Octahedral Complexes, Potential Surfaces, Electric Dipoles

INDEX

November 1980, 144 pp., illus., 5½" × 8¼"
Hardbound 30452 $14.50

*Benjamin/Cummings Title

Price is subject to change without notice. Outside U.S.A. price listed may vary.

Addison-Wesley Publishing Company, Inc.
Advanced Book Program
Reading, Massachusetts 01867, U.S.A.

6 Byfield Street, North Ryde, New South Wales 2113, Australia
53 Bedford Square, London WC1B 3DZ, U.K
De Lairessestr. 90, Amsterdam 1071, The Netherlands
36 Prince Andrew Pl., Don Mills, Ontario M3C 2T8, Canada

Circle No.39 on Reader Enquiry Card.

to chapters on muscle and nerve cells.

Integration of cellular and molecular matters could be improved in places; for example, discussion of self-assembly is divorced from some of its best examples, the G/F transitions of filaments and tubules, and there seems to be no hint of the possible role of glycosylation in intracellular traffic. There is a moderate if uneven emphasis on experimental evidence. For example, we meet puromycin in relation to protein synthesis, but not EGTA in excitation-contraction coupling. Though X-ray diffraction appears amongst instrumental methods, the task of explaining the importance of low-angle diffraction by intact muscle is avoided. Further there seem to be rather too many small oddities, even for a volume with so wide a sweep: EDTA works in one experiment by "binding to phosphate groups"; freeze fracture may "disclose the outer or inner surface of a membrane or may even split the membrane lengthwise". Worst, in context of cancer "normal cultured cells generally divide every 24 hours as long as they float freely in the medium". This last points to a wider weakness: of the two main tools of cell biology today, the electron microscope and cell-culture, the authors are clearly more at home with the former. Otherwise they would not propagate the myth of contact inhibition of mitosis, subsume the locomotion of tissue cells with that of amoebae, or omit all but the vaguest reference to fibronectin. The indexing could usefully be three times denser.

Nevertheless, though this book looks more successfully down the hierarchy than up, it does show that teaching centred on the cell (rather than the animal or enzyme) need not lack coherence or excitement. For such a course it would be useful but far from definitive. □

John Edwards is Senior Lecturer in the Department of Cell Biology at the University of Glasgow.

Textbook evolution

K. Murray

DNA Replication. By Arthur Kornberg. Pp.724. (W.H. Freeman: 1980.) £19.20, $34.95.

If one regards *DNA Replication* as a second edition of *DNA Synthesis,* one should read the two side by side. One then sees a remarkably concise summary of the advances in nucleic acid enzymology that have occurred in the five years that separate the two volumes, and a lucid integration of the role and interrelations of several newly discovered enzymes in the complex and intricate processes whereby genetic information is maintained and utilized. The evolution of *DNA Replication* from *DNA Synthesis* has involved a number of deletions, an appreciable amount of rearrangement and, not surprisingly since it is nearly twice the size of its progenitor, numerous insertions that confer a clear advantage upon the descendant.

The introductory chapters on the structure and functions of DNA and the biosynthesis of DNA precursors are expanded appreciably, while the chapters on DNA polymerase I of *E. coli* and other prokaryotic DNA polymerases remain largely as the classic accounts that they were, which I believe should still be read by all students of nucleic acid biochemistry — incidentally I am glad that one of my favourite quotations of Kornberg remains: "Failure to detect an enzyme activity in a cell-free extract means only that." (p.160).

It is in the discussion of eukaryotic DNA polymerases in Chapter 6 that the expansion and reorganization of material becomes really apparent. The many enzymes involved in various phases of DNA replication are now allotted individual chapters and this, in my opinion, improves the subsequent discussion of replication mechanisms and their regulation. The chapter on bacterial DNA viruses is appropriately updated and expanded, and animal viruses are similarly treated in a new chapter of their own. As before, chapters on DNA repair, recombination and restriction (and now transformation also) and on the synthesis of genes and chromosomes conclude the book. The final chapter is completely revised and introduces modern advances in DNA sequence determination and recombination DNA technology, topics that merit a book of their own. It is, as one expected, an excellent book written with clarity and an economy for transferring information which makes small bacterial viruses look extravagant, and with the authority of one who has done much to lay the foundations of the subject and set standards for others to emulate.

Like *DNA Synthesis, DNA Replication* will be widely valued both as text and as a reference book. It should be on the shopping list of all students in biological subjects — and chemists too, for that matter, in this era of biotechnology — but it will be equally valuable to research workers. If it stimulates the need for a third version, I am sure Kornberg will feel suitably rewarded. □

K. Murray is at the European Molecular Biology Laboratory, Heidelberg.

Biomembranes — agreement and discord

Dennis Chapman

Biological Membranes: Their Structure and Function. 2nd Edn. By R. Harrison and G. G. Lunt. Pp.288. (Blackie/Halsted: 1980.) Hbk £17.95; pbk £7.95, $21.95.
Introduction to Biological Membranes. By M. K. Jain and R. C. Wagner. Pp.382. (Wiley-Interscience: 1980.) £18.40, $43.25.

The ways in which a scientific consensus is reached on a subject must always be fascinating to philosophers of science. The consensus view on biomembrane structure is often expressed in terms of the fluid-mosaic model. At the same time it is appreciated by all the workers in the field that biomembranes exist which are not fluid but rigid, where the proteins are fixed in position, not mobile, and where a mosaic structure does not occur — instead, the proteins are arranged in a crystalline manner.

These topics and the relationship between structure and function are discussed in these two books. Both are up to date and deal with broadly the same area, covering electron microscopy, membrane preparations, structural organization, transport and receptor functions. The book by Jain and Wagner is perhaps the more elegantly written, and contains good diagrams with clear summaries and useful references, but both books will be useful to graduate and undergraduate students interested in having a concise review of the present situation.

Given two such similar books, it was interesting to see how they discussed an area of research which has excited interest over the last two or three years. This is the topic of protein–lipid interactions in biomembranes. Jain and Wagner comment

> . . . various lipid activated enzymes contain an annulus of tightly bound lipid molecules . . . . Such lipid has distinctive properties, cannot be readily extracted, is *not* readily exchangeable with the more fluid lipid . . . .

Harrison and Lunt comment

> . . . a number of membrane-bound enzymes have been shown to be surrounded by an annulus of immobilized boundary lipid, the local phase transitions of which are reflected in the activity of the enzyme . . . and boundary lipid does not preclude *rapid exchange* with surrounding lipid molecules.

Perhaps in the next editions a new consensus view on boundary lipids will be reached! However these quotations merely show how difficult it is to give a snapshot impression of a scientific area which is still alive and active. □

Dennis Chapman is Professor of Biophysical Chemistry at the Royal Free Hospital School of Medicine, University of London.

Introductory chemistry in the market place

Arden P. Zipp

Fundamentals of Chemistry. 4th Edn. By F. Brescia, J. Arents, H. Meislich and A. Turk. Pp.781. (Academic: 1980.) $21.95, £13. *Foundations of College Chemistry.* 3rd Edn. By D. B. Murphy and V. Rousseau. Pp. 767. (Wiley: 1980.) $29.20, £12.45. *Basic Chemistry: General, Organic, Biological.* By D.M. Callewaert and J. Genyea. Pp.837 (Worth: 1980.) $19.95, £9.25.

THE teaching of introductory chemistry has given rise to much discussion in recent years due to two main factors. The first of these concerns the question of the topics to be taught and the level of presentation to be attempted. The second has to do with the wide variation in the abilities of the students who take such courses. These range from individuals who have received good preparation in secondary school and have little difficulty dealing with the abstract concepts which are inherent in the subject matter, to students who are ill-prepared to deal with these concepts. Although the second factor may be more common to US colleges and universities, it would be surprising if it were totally absent from classes in other countries as well.

In order to compete in the current market, textbook authors are called upon to provide books which are both broad in scope and flexible in level. The three discussed here are intended to accommodate students of varying backgrounds with potentially different career goals who need a general chemistry course; that is, students who are likely to major in science although not necessarily in chemistry. Despite this common objective, and a number of similarities among them, these books actually provide a wide choice in level and approach, and each offers something unique.

Each book is about 800 pages in length, has been produced in an attractive manner with an extensive use of colour, and all seem to have been well-edited with few errors. All three cover the topics usually treated in general chemistry: stoichiometry, atomic structure, bonding, states of matter, thermodynamics, equilibrium, kinetics, acids and bases, electrochemistry, elementary organic chemistry, polymers and (except for Callewaert and Genyea) coordination and nuclear chemistry.

Of the three, the book by Brescia *et al.* seems to be the most student-orientated, each chapter containing learning objectives, key terms, self-tests, worked-out examples and many practice problems, with answers given in an appendix. (A companion student guide is also available for this text.) Each of the other books possesses only some of these features. The authors have adopted the technique of presenting topics more than once, first on a qualitative level, then on a more intensive basis. Although this approach is successful in most cases, it is surprising to find the second stage of bonding deferred to Chapter 20 after the initial treatment has been presented in Chapter 9. The bulk of thermodynamics is presented as it relates to phase changes, rather than in a more abstract fashion. Descriptive chemistry appears mostly as a means of illustrating principles rather than in a separate chapter or chapters. A chapter dealing with environmental chemistry is presented, as is one on radiation and matter. The latter provides coverage of infrared, ultraviolet, nuclear magnetic resonance and mass spectroscopy in much greater depth than is usually encountered at this level, although some instructors may find it appropriate.

The text by Murphy and Rousseau is the highest in level mathematically and provides the greatest emphasis on physical chemistry. It presents numerous derivations including the Bohr radius, van't Hoff and Balmer equations, and the gas laws (from kinetic theory). Three centre bonds (boron hydrides) are discussed in some detail and the Debye–Huckel theory is mentioned. It is the only one of the three books in which descriptive chemistry is discussed separately. A 35-page chapter on group IV (emphasizing the variations in properties observed as the family is descended) precedes a chapter on organic chemistry, which is followed in turn by one of 57 pages in which the remaining non-metals are discussed. A separate chapter on metallurgy is also presented. The book offers many problems, both worked-out in the text and at the ends of chapters (with answers to odd-numbered ones given in an appendix), and provides a list of suggestions for further reading after each chapter. In addition, a student guide and solutions manual are available.

The distinguishing feature of the Callewaert and Genyea text is the heavy emphasis on organic and biological chemistry. Approximately one-half of the book is devoted to "general" chemistry with the remainder being divided about evenly into three sections devoted to organic molecules, biological molecules and metabolism. In addition, many of the examples used in the first portion of the text are biochemical in nature. This text was the hardest one for me to evaluate, having had little experience teaching such a course. Although the authors say that their intent is "explaining essential chemical concepts rather than providing a cursory treatment of a much larger number of topics", some of the discussions seem too brief to provide a real understanding of the material. For example, kinetic molecular theory, solids, liquids, gases and elementary thermodynamics are discussed in a chapter of 27 pages. Similarly, equilibrium and kinetics (including the disproven direct collision mechanism for the H_2–I_2 reaction) are "covered" in 21 pages. Further, it is surprising to find electron configurations relegated to an optional section of a chapter and to encounter a discussion of bonding without substantial use of orbitals. This approach may be appropriate and even necessary in order to cover all the material included in the text (and presumably in such courses), although the book's use in the broad general chemistry market is likely to be limited. The book does give study objectives, useful chapter summaries and solutions to exercises at the end of each chapter, but has relatively few problems and worked-out examples and is the least mathematical of the three.

These three texts represent only a small sample of the total number which is currently available, although they do provide some indication of the almost limitless variety which exists in this highly competitive market. ☐

Arden P. Zipp is Professor of Chemistry at the State University of New York, Cortland, New York.

Biophysical challenge

R. G. Gosling

Topics in Classical Biophysics. By H. J. Metcalf. Pp.286. (Prentice-Hall: 1980.) Pbk £6.45, $9.95.

MY readiness to review this book lay in its intriguing title and my need of a broadly based text for students pursuing a BSc degree in "physics with medical applications". The format is primarily that of a textbook, although the author suggests it "is not intended for students alone". The first eight of its nine chapters present appropriate basic theory — biomechanics, heat and energy, fluids, blood circulation, feedback and control, nerve cells, sound and hearing, light, colour and vision. The final chapter is a somewhat mixed bag entitled "Experimental Techniques" but, with its emphasis firmly on the use of physics to acquire information, nicely draws together the joint application of various principles previously discussed.

Essentially, the theoretical treatment is that of classical continuum physics, with basic equations developed at the beginning of each chapter and then applied to some selected topics appropriate to that field. Whilst the question as to whether this justifies the title *Classical Biophysics* could be debated, I found that the material chosen represented a useful introduction for any undergraduate interested in the applications of physics to medicine. The only pre-requirement is an elementary knowledge of calculus, for example additional maths at "O" level in the UK or an introductory year college math course in the USA.

Because of the wide range of natural phenomena discussed, the theory that can be developed in the commendably brief compass of 275 pages is necessarily kept at a lowish level. The penalty is that arguments are sometimes terminated a little prematurely. For example, in Chapter 3 on fluids, the forces acting in a Newtonian liquid flowing in a rigid tube are clearly deduced, but then used only in the special case of a constant pressure gradient. This precludes any application of such theory in discussing arterial haemodynamics in the next chapter. However, it does provide a clear "launching point" for any lecturer using the book as a class text.

It is my impression that this book will present first- and second-year students of physics, biology or medicine with challenging reading and I look forward to using it with such classes next year. It may prove a useful introduction to any courses concerned with the application of classical physics to biomedical problems. ☐

R. G. Gosling is University Reader in Physics Applied to Medicine at Guy's Hospital Medical School, University of London.

Food — science and politics

J.W.T. Dickerson

Nutrition: in Perspective. By Patricia A. Kreutler. Pp.700. (Prentice-Hall: 1980.) £12.30, $18.95. Workbook also available.

THERE has recently been something of a spate of books on nutrition, and new ones must therefore face increased competition. The market itself has also increased, for we are becoming much more aware of the relevance of food to life and health in the modern world. Moreover, all shades of interest and competence, from the lay reader to the academic, can profit from the study of the subject. This book has been aimed somewhere in the middle of the range. It "has been written to provide students with an understanding of the science of nutrition"; to acquaint them "with the issues facing nutritionists, scientific and government leaders, and consumers in contemporary society", and "to enable students to translate knowledge into practice".

The first part of the book is concerned with the scientific principles of nutrition, and contains nine chapters. The second part discusses aspects of nutrition for everyday living, in seven chapters. The author and publishers are to be congratulated on the plan and layout of the book. The text is clear, direct and liberally illustrated throughout with tables, figures, photographs and the occasional cartoon. The author is clearly an expert communicator of her subject. References in the text have been kept to a minimum and are listed, with recommendations for additional reading, at the end of each chapter.

A feature of the book is the introduction, as appropriate, of "Perspectives" — applications to health or topics of contemporary concern. Thus, in the chapter on carbohydrates, the perspectives are sugars in foods, diabetes and hypoglycaemia, and dietary fibre. The chapter on lipids contains a perspective on diet and heart disease. Perspectives on sodium and hypertension, iron enrichment and the politics of fluoridation are inserted in the chapter on minerals, whereas fad diets are discussed in the chapter on food in contemporary society.

The whole book is a pleasure to read. It is interesting, informative and contains a great deal of valuable information. ☐

J.W.T. Dickerson is Professor of Human Nutrition at the University of Surrey.

Chemistry for life

Peter Garratt

Essential Organic Chemistry for Students of the Life Sciences. By A.P. Ryles, K. Smith and R.S. Ward. Pp.306. (Wiley: 1980.) Hbk £14.90, $44.10; pbk £5.25, $15.75. *Organic Chemistry: The Basis of Life.* By Bernard Miller. Pp.472. (Benjamin/Cummings/Addison-Wesley: 1980.) $18.95, £11.85.

THE problem of teaching chemistry to students in the biological sciences is one of striking a balance between relevance and comprehension. The student, who probably has no great interest or aptitude for chemistry, wants only that material which is immediately relevant to his biological studies, whereas the teacher knows that to have any appreciation of the role of chemistry in biology the student must understand the concepts and principles on which the subject is based. These two books, although emanating from different continents and very different educational systems, clearly illustrate that this problem is also felt by writers of texts for such students.

Both books deal, as their titles imply, only with the organic component of the chemistry required by biologists. RSW (for short) has three introductory chapters, the first on bonding, functional groups and mechanism, the second on stereochemistry and the third on techniques. Then follow nine chapters presenting a conventional and systematic treatment of organic chemistry from the alkanes to the proteins and nucleic acids. The last three chapters deal with tetrapyrroles, physiologically important compounds and metabolism and biosynthesis. Miller has an introductory chapter in which "basic principles" are sketchily and sometimes erroneously treated, followed by a chapter on the alkanes whch includes further introductory material on bonding. The remaining 16 chapters give a similar outline of organic chemistry to that found in RSW, except that the ordering of material differs. For example, chirality is introduced on p.23 in RSW but not until p.197 in Miller, rather late for a book subtitled the *Basis of Life.* Miller integrates the biological examples more intimately with the chemistry than do the British authors, and each chapter ends with a summary of the important principles and concepts found therein. Both texts have problems distributed within the chapters, and there is a student study guide available for Miller and one promised for RSW.

The layout, appearance and style of a text are important to its student user. Both texts are well produced, Miller being a two-colour printing, in which the colour has been judiciously used, while RSW is in monochrome. Miller has clear structural diagrams and well drawn, and therefore useful, representations of space-filling models, but it also has unnecessary photographs of oil rigs and similar irrelevant objects. RSW has, in general, clear structural formulae except for those representing the conformations of cyclohexane and for ascorbic acid, the nicotinamide co-enzymes and co-enzyme A. RSW also uses Haworth rather than conformational formulae for the carbohydrates. Since shape is so important for biological activity, conformational structures are much more instructive, and the treatment in Miller, which uses such structures, is clearly superior.

Both of these texts would form a satisfactory basis for a course in organic chemistry for biologists, but whether much would be gained over using a standard chemistry text such as Solomons' *Organic Chemistry* (Wiley, 2nd Edn 1980 — see over for review) is a moot point. Miller certainly makes a more direct attempt to show the relevance of the chemistry to biological molecules and processes than does RSW and the style is easier to read, but the sections on orbitals and bonding would require remedial attention from the lecturer. ☐

Peter Garratt, a Reader in the Department of Chemistry at University College, University of London, regularly lectures to biology students on chemistry.

Organic chemistry, in theory . . .

John Mann

Organic Chemistry. 2nd Edn. By T. W. G. Solomons. Pp.1066. (Wiley: 1980.) Hbk £15.25, $33.25; pbk £7.90, $18. *Organic Chemistry.* By D. S. Kemp and F. Vellaccio. Pp.1421. (Worth: 1980.) Hbk $26.95, £10.25; pbk $14.95, £5.

It is not often that an organic chemistry textbook is the subject of a plagiarism suit, but such was the fate that befell the first edition of *Organic Chemistry* by T. W. G. Solomons. After a four-month hearing the book and its author were cleared in March 1980, just prior to the release of the second edition. Personally, I have always found the other book in question (*Organic Chemistry* by R. T. Morrison and R. N. Boyd; Allyn & Bacon, 3rd Edn 1973) pretty indigestible and rather old-fashioned, though students seem to like the summaries of functional group chemistry.

Professor Solomons's book has similar summaries at the end of most chapters, but it also has a great many features which enhance its value relative to Morrison and Boyd. The order of treatment is conventional, commencing with hydrocarbons and bonding, and proceeding through the various functional groups to carbohydrates and proteins. Well-chosen special topics are interspersed throughout, and complement the chapters in which they appear. We thus have naturally-occurring alkenes and the photochemistry of vision at the end of the chapter on conjugated unsaturated systems; a beautifully illustrated and lucid account of protein biosynthesis accompanies the chapter on amino acids and proteins; and there are also useful sections on applications of the Woodward–Hoffmann rules, halo-organic insecticides, polymerization, and biosynthesis of fatty acids and steroids. One small puzzle is why thallation deserves a six-page section of its own, while the Wittig reaction and organosulphur chemistry (both much more important in modern organic chemistry) receive only scant attention. The chapter on spectroscopic techniques, now a must for any organic text, provides a very good introduction to the subject, and is comparable with the corresponding section in *Introduction to Organic Chemistry* by A. Streitweiser and C. H. Heathcock (Macmillan, New York, 1976), the book that we recommend to our students.

There are a few disappointments in Professor Solomons's book. For example the absence of a section on the strategy of synthesis, and the rather old-fashioned account of synthesis and reactions of β-dicarbonyl compounds which contains a long list of named reactions, but little in the way of modern synthetic applications or mention of their relevance to biochemistry. These criticisms aside, the main functional group chemistry is well done, the book is beautifully produced with two-tone drawings and section headings, and there are clear chemical structures throughout. It is readable and reasonably up-to-date in most areas, and has innumerable useful problems throughout, with answers supplied at the end of the book or in the separate solutions manual.

The authors of the second glossy American book under review, D. S. Kemp and F. Vellaccio, claim that there is a need for a text emphasizing the chemistry of the oxygen- and nitrogen-containing functional groups. The book is designed for an American one-year/three-semester course (as I suspect Solomons's book is), and it is divided into three main parts: chemistry of functional groups and structure proof; chemistry of hydrocarbons and carbonyl chemistry; and advanced topics. The order of treatment is thus somewhat different from that of most other texts. Apart from this difference, the treatment of topics is rather similar to that in Solomons, except that there is more emphasis of applications to organic synthesis. The authors claim that their treatment of carbonyl chemistry is the "most accessible and interesting of any that are available", and in this they are probably correct. The chapter on spectroscopic methods provides a good introduction to NMR and IR, but is very weak on mass spectrometry (a mere three and a half pages), while chapters on carbohydrates, proteins and nucleic acids are good, though not so well done as in Solomons. There are good sections on the Woodward–Hoffmann rules; rearrangement reactions; reactive intermediates; and organometallic reagents, which includes as special topics hydroboration, silanes, organomercury compounds and thallation. Coverage of heterocyclic chemistry is, however, rather weak, and natural products are dealt with in a somewhat fragmented way. The penultimate chapter, on the synthesis of six compounds of theoretical or practical importance, is novel and makes interesting reading. It includes, amongst other topics the conversion of penicillins into cephalosporins, and capped porphyrins. Again, there are useful problems throughout, and the 670-page workbook and solutions manual should find many friends amongst both students and academics. Throughout, the illustrations are very good, but the book is not quite so lucid or readable as Solomons, and many students will miss the reaction summaries.

Both books are modern texts comparable to the excellent Streitweiser and Heathcock, and will I am sure sell well, especially in the USA. However neither of them has everything quite right, and the ultimate text has still to be written. But then, there is probably no such thing! □

John Mann is a Lecturer in Organic Chemistry at the University of Reading.

. . . and in practice

Paul Schatz

Experimental Organic Chemistry. By Michael P. Doyle and William S. Mungall. Pp.490. (Wiley: 1980.) Hbk £10.10, $23.70; pbk £5.95, $13. *The Systematic Identification of Organic Compounds.* 6th Edn. By R.L. Shriner *et al.* Pp.604. (Wiley: 1980.) Hbk £13, $30.55; pbk £6.90, $15.90.

The pedagogical aim of an organic laboratory course is to introduce the student to the basic techniques used by chemists for the preparation, separation, purification and identification of organic compounds. Besides demonstrating such techniques as distillation, recrystallization, chromatography and extraction, the laboratory is a showplace for demonstrating the chemical properties of the various organic functional groups. Most introductory laboratory courses use what might be dubbed "the synthetic approach". The techniques are introduced in the context of carrying out typical organic reactions, such as an Sn1 substitution or an E2 elimination, isolating and purifying the product, and confirming the identity of the product. An alternative approach to teaching in the organic chemistry laboratory could be called "the characterization approach". In this approach, the laboratory techniques are taught in the context of classifying and identifying organic compounds by functional groups. Typically, derivatives of compounds are prepared and purified, and then identified by measuring physical properties. Although most qualitative analysis courses are usually reserved for intermediate or advanced level undergraduate students, many modern introductory courses employing the "synthetic approach" also incorporate a qualitative analysis component into the curriculum.

There are a large number of textbooks (or laboratory manuals) available for the introductory laboratory and the manner in which they treat the material is very similar. There is a section introducing basic techniques, there is a section on qualitative analysis and there is a section containing a variety of experiments for the instructor (or student) to choose from. *Experimental Organic Chemistry* follows a similar pattern. What distinguishes it from other texts are the experiments. They are not a rehash of familiar experiments, but a set of new ones. Moreover, they are interrelated since the product of one experiment can serve as the starting material for another. The authors suggest that besides being educationally beneficial, it can also whittle down the cost of the laboratory course. By choosing readily available chemicals for both the starting materials and the products, the authors have avoided the pitfall of locking the user into a prescribed set of experiments and the vexing problem of what to do with the student who gets a poor yield. The variety of experiments is impressive — Sn1 substitution, Sn2 substitution, enzymatic hydrolysis, phase transfer catalysis and so on — and they include a good collection of short multistep (two or three steps) syntheses. The book has much to recommend it besides the experiments. The techniques are explained fully and clearly, the experimental procedures are detailed and clear, the artwork is accurate and well drawn, and the reproductions of the IR and NMR spectra are sharp. All in all, this is a well crafted book, definitely one to be considered when choosing a text for laboratory courses.

One of the most familiar books for the organic qualitative analysis course has been *The Systematic Identification of Organic Compounds.* A sixth edition has been recently published and this edition is markedly different from its predecessors. The most obvious difference is the size; the newest edition is approximately 50% wider than the fifth edition. This larger size is needed to contain the expanded text. For the most part, all of the elements of the older editions are incorporated in this edition, including the measurement of physical properties, the classification tests, the preparation of derivatives and the comprehensive tables of derivatives. What has been added are more extensive discussions of spectroscopic methods of identification and chromatographic methods of separation. The latter topic was completely omitted in previous editions.

Other additions are a larger introduction, including a discussion of laboratory safety, a more detailed explanation of basic laboratory techniques such as distillation and recrystallization, and a chapter on the literature of organic chemistry (principally an annotated bibliography). The book would be useful for the classic organic qualitative analysis course (if any of these still exist). It could also be used in an introductory organic laboratory course emphasizing qualitative analysis if supplementary material describing the theory of the basic techniques is available, since no theory is explained in the text. □

Paul Schatz is Director of Organic Laboratories at the University of Wisconsin-Madison.

Joys of general chemistry

Charles A. McAuliffe

Chemistry (With Inorganic Qualitative Analysis). By T. Moeller *et al.* Pp.1085. (Academic: 1980.) £14, $21.95. *Chemistry: A Systematic Approach.* By Harry H. Sisler, Richard D. Dresdner and William T. Mooney Jr. Pp.960. (Oxford University Press: 1980.) £13.50, $19.95. *First Year Chemistry.* By J. M. Coxon, J. E. Fergusson and L. F. Phillips. Pp.376. (Edward Arnold: 1980.) Pbk £9.75.

I HAVE been reviewing freshman chemistry texts for this edition of *Nature* for four years and, despite a long association with a number of American universities and their methods of teaching, each particular year of carrying out this task offers me an opportunity of seeing new facets in these books. For example, the best of this type of book, in addition to expounding chemical fact and theory, also stresses the application of chemistry to everyday life. I have been particularly pleased by attempts in a number of excellent texts at the way in which environmental problems are related to the chemistry being illustrated. For example, consideration of the reactions involved in the generation of photochemical smog cannot only make a student appreciate the chemistry of the nitrogen oxides and ozone, but also that of organic free radicals and the importance of light in initiating chemical reactions. Similarly, dealing with elemental cycles in the biosphere not only emphasizes their importance to life but also serves to make simple chemistry very real; for example, the most important of all chemical processes, photosynthesis, should not be dealt with in isolation, large subject though it is, but can in a freshman text be linked to the oxygen and carbon cycles and fossil fuel generation.

The enormous market for freshman (or general chemistry) texts in the United States ensures that many are available. In previous years I have emphasized how awful some of these texts are, and one shudders to think of them being foisted on students by their authors; many potentially creative chemists must be lost in this way. However, the three texts dealt with here are, in their different ways, a joy to behold.

Chemistry (with Inorganic Qualitative Analysis) has so much to recommend it. There are 32 chapters and it is sensible to give a brief indication of their range so that one can understand the scope of a large textbook such as this. Atomic structure and nuclear chemistry, chemical bonding, periodic properties, thermodynamics, kinetics and electrochemistry provide an adequate theoretical basis. Descriptive group chemistry and some, though never enough, organic chemistry is related to reactivity and stoichiometry. There are several chapters on qualitative analysis.

It is difficult to single out individual sections of this book for praise. I used the extensive index a great deal in examining the book. Nothing which should be included has been omitted; the text is concise and illuminating. Like many such books, the diagrams are excellent. I particularly enjoyed the copious vignettes entitled "Thoughts on Chemistry". These are enlightening quotes, a page or so in length, which range from "The Requisites of a Good Hypothesis" (Robert Boyle), "The World's Biggest Membrane" (from Lewis Thomas's memorable observations in *The Lives of a Cell: Notes of a Biology Watcher*), "The Chemical History of a Candle" (Michael Faraday), "The Centrality of Chemistry" (Martin Sherwood), to "The Epigrams of Remigius Fresenious". This book is to be highly recommended. I have only two criticisms. The minor criticism is that on the same page (p. 21) which shows the structure of chlorophyll *a* there is advice on how to pronounce $NaNO_3$ ("N-A-N-oh-three") and H_2SO_4 ("H-two-S-oh-four") — this grated on me. More serious is the large amount of space they devote to inorganic qualitative analysis. This is a grindingly boring topic and, if included at all, should be kept to a minimum. The authors would do well to extend their good descriptions of infrared, ultraviolet and nuclear magnetic resonance spectroscopy.

Chemistry: A Systematic Approach is also to be recommended. It is comprehensive and I particularly enjoyed the chapter on the chemistry of living systems. An instructor's manual (by W. H. Myers) is available.

First Year Chemistry is a somewhat different book. The authors aim to produce a text in which approximately equal weight is given to physical, inorganic and organic chemistry, and minimizes the divisions. It is a very good attempt, but the book may have limited appeal. It assumes a much greater basic knowledge than do the other texts and may be only suitable for the New Zealand system, in which the authors work.

Finally, might I appeal to all of these authors? Expressing photosynthesis as

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2 + \text{energy}$$

is very misleading, even though this is the way major biology texts frequently express it. Carbon dioxide does not react with water as implied by the equation. Further explanation would be enlightening □

Charles A. McAuliffe is Reader in Chemistry at UMIST and Adjunct Professor of Chemistry at Auburn University, Alabama.

Quantum mechanics for chemists

A. J. Stone

Orbitals and Symmetry. By D. S. Urch. Pp.256. (Macmillan: London, 1979.) Pbk £5.95. *Chemical Structure and Bonding.* By R. L. DeKock and H. B. Gray. Pp.491. (Benjamin/Cummings/Addison-Wesley: 1980.) $21.95, £9.95.

THE flood of textbooks on quantum mechanics for chemists puts one in mind of that wilderness of monkeys which are supposed in time to produce the works of Shakespeare — not because the results are similar but because we are prompted to ask: why bother? We already have textbooks at every level which give clear, sound coverage of the principles of the subject. It is not enough for a new book to do the same; it must provide new insights, take account of advances in experimental and theoretical technique, or treat some aspect of the subject substantially better than before.

DeKock and Gray's angle is to relate the discussion of bonding and structure to the results of photoelectron spectroscopy. However, the book begins with the Bohr theory (what other superannuated theory has such resilience that it is still being taught 50 years after it was superseded?) and there is much emphasis on Lewis octets and valence-shell–electron-pair repulsions. Valence-bond theory is discussed at length, and molecular-orbital theory takes a bad second place. For example, a superficial and rather inaccurate account of a simple Walsh diagram is followed by the statement that "predicting molecular shape using molecular-orbital theory. . . is much more tedious than the valence-shell electron-pair repulsion method". This attitude to molecular orbital theory is curious in view of the detailed account of the relationship between photoelectron spectroscopy and bonding, which is the one useful feature of the book and which is done rather well — indeed, at a level which is noticeably higher than the rest of the book.

Urch's book is a reprint of a volume originally published in 1970 in the abortive Penguin Library of Physical Sciences, and although the preface implies some revision of the text, this version looks virtually identical to the original. Some errors have been corrected, but others have not, and the bibliography remains unchanged. My main criticism is that there is not enough attention given to the business of showing the student how to do the calculations. For example, the representation matrices for a set of d orbitals in the octahedral group (a notorious source of difficulty) are simply written down without comment. More importantly, the procedure for obtaining symmetry orbitals (the projection formula) is never mentioned, and symmetry orbitals are blandly produced, not merely without explanation but without any apparent appreciation that the reader might not see where they came from. The text could quite easily have been revised to correct these failings, but the opportunity has been lost and a potentially useful account remains incomplete.

A. J. Stone is a Lecturer in Theoretical Chemistry at the University of Cambridge.

Environmental chemistry, theory and practice

Roy M. Harrison

Environmental Chemistry: The Earth–Air–Water Factory. By R. W. Raiswell *et al.* Pp.184. (Edward Arnold/Halsted: 1980.) Flexi £4.95, $14.95. *Experiments in Environmental Chemistry: A Laboratory Manual.* By P. D. Vowles and D. W. Connell. Pp.102. (Pergamon: 1980.) Hbk £10, $23; pbk £4.50, $9.95.

THE teaching of environmental chemistry as a component of environmental science, ecology and chemistry courses at university level has developed and expanded considerably over recent years, largely in the absence of adequate textbooks. A number of books have appeared recently, but there still exists a need for sound teaching texts.

Environmental Chemistry: The Earth–Air–Water Factory provides a broad introduction to the field written mainly from the viewpoint of the geochemist. The factory analogy is used apparently in an attempt to provide a better integration of subject matter and to encourage the student's appreciation of the strong links between the various areas of environmental chemistry. The book assumes rather little chemical background and even includes a glossary of chemical terms. It is written with clarity by authors acknowledged as authorities in their respective fields. The great detraction, however, is that no text of 154 pages (excluding glossary and index) can more than touch the surface of this vast subject area. It is difficult to see the book being used for university teaching other than at the most basic level as it will satisfy neither the lecturer's aims nor the student's curiosity.

Experiments in Environmental Chemistry: A Laboratory Manual concentrates upon pollution and ecological chemistry, the majority of experiments being analytical in nature. The presentation is similar to that used in laboratory manuals in universities, with an introductory section followed by experimental instructions and then guidance for writing up. The factual material is occasionally unreliable, although in some instances this arises from typographical errors. The experiments require a fair variety of analytical equipment which may be beyond the resources of some teaching laboratories, and some demand rather problematic starting materials — human milk; butterfly larvae and plants from the Southern Hemisphere; replicate 0.2 g samples of household paint removed from surfaces, for example. The book contains some excellent ideas for experiments but is more likely to find use as inspiration for the teacher than as a laboratory manual used directly by the student. □

Roy M. Harrison is a Lecturer in the Department of Environmental Sciences at the University of Lancaster.

Environmental what?

John Barkham

Ecology and Environmental Management. By C. C. Park. Pp. 272. (Butterworths/Westview, Boulder, Colorado: 1980.) £12.50, $26.25. *Natural Resource Conservation: An Ecological Approach.* 3rd Edn. By Oliver S. Owen. Pp.883. (Macmillan, New York/Collier Macmillan, London: 1980.) £12.95, $19.95. *Introduction to Environmental Science.* By J.M. Moran, M.D. Morgan and J.H. Wiersma. Pp.658. (W.H. Freeman: 1980.) £19.95, £9.80. Instructor's guide also available.

MANY different disciplines contribute to attempts to understand the environment. The rapid development of courses in environmental sciences or studies in recent years reflects the need to look beyond geography, geology, biology or economics alone for the range of material required. However, the variety of university and college courses calling themselves "environmental" suggests widely differing views as to what are the key subject areas, concepts and techniques to which students should be addressing themselves. That there are degree courses in "environmental *science*" implies that there are those bold or unwise enough to define a distinct discipline. A more sensible and pragmatic approach is to regard the sciences concerned with the environment as a federation. They may usefully be gathered together in one department so that environmental problems may be explored from a number of different perspectives. This flexible view allows students to adopt an eclectic approach to learning, drawing upon sources of information which appear to be useful whatever their pedigree may be. So it is likely that books written about environmental sciences from a variety of points of view will be useful to different students at different times.

The three books considered here all examine problems of environmental management from an ecological perspective. They appear to be directed towards first-level university and college courses. Questions of management always stimulate student interest and passion, but the soundness with which undergraduates approach such issues usually depends upon the adequacy of their grounding in basic biological, environmental and social sciences. All three books start from the assumption that their readers know little or no ecology and attempt to give them a grounding in it. Park's book, *Ecology and Environmental Management*, devotes about 170 pages to this material, about two-thirds of the book, although the first 20 pages are spent in largely sterile justification of the notion that geography is "centrally placed to contribute effectively to the solution of many environmental problems". This section rather underlines the fact that this is an ecological text written by a geographer for geographers and environmental scientists. Ecologists will find the balance rather odd and it is by no means clear at times why, in relation to environmental management, we should be learning the details of some ecological processes. There are some extraordinary statements, too, including a comment on the ". . . lack of interest within ecology for animal populations . . ." (p.39). For me, there is a definitely dated feel to the first two-thirds of this book, which is perhaps emphasized by the predominance of geographical references and the paucity of ecological ones from the 1970s.

In Owen's book, *Natural Resource Conservation*, the ecological basis provided is the barest minimum necessary and is rather scattered about within the text following a short section near the beginning. Moran, Morgan and Wiersma's *Introduction to Environmental Science* is the best balanced from this point of view, although the title is entirely misleading. Ecological fundamentals occupy about as much space as in Park's work, the approach is traditional — but none the worse for that — and, above all, the material is well organized. It is obviously a carefully planned book and I wonder whether this is the pay-off for good teamwork. There are conclusions, useful summary statements, review questions and boxes of information where deeper explanation of a topic is felt necessary.

In the environmental management sections, Owen's and Moran *et al.'s* American texts are very different from Park's English approach. Having waded through Owen I felt blasted with the sheer volume of material on every kind of resource and pollution problem. No question of a balanced view here. I felt the author wanted the reader to join him in the corral of the good boys. However, there is a lot of qualitative information about every kind of problem, solutions are discussed and the work is liberally illustrated. Moran *et al.* cover the same ground in about the same amount of space but in a well-ordered, less frenetic manner and with fewer gee-whiz statistics. Both books use very largely American case studies and examples and refer to American ways and means of tackling these problems.

Park's book examines resource, pollution and management issues relevant to Britain. Only two pages are devoted to problems of a global scale. Students concerned with nature conservation will find a lot of useful material, particularly in the tables. Others taking a broad-based course will get a reasonable overview of the subject.

In terms of a best buy, Owen's book may well appeal to a section of the American market, but not in Britain. Park's will have a limited appeal mainly to geographers in Britain, but not in America. Moran *et al.* is best value for money and may sell to a limited extent in Britain as well as in America, although a title such as *An Ecological Approach to Environmental Management* would be more appropriate. Sadly, none of these books succeed in giving the student a clear global perspective which is not at the same time banal. It is also clear to me that any student who is exposed only to the approach adopted here will make a disastrous environmental manager. □

John Barkham is in the School of Environmental Sciences, University of East Anglia.

Views of the landscape

C. Kidson

Process in Geomorphology. Edited by C. Embleton and J. Thornes. Pp.436. (Edward Arnold/Halsted: 1979.) Hbk £16, $44.95; pbk £8.95, $19.95. *Geomorphological Processes.* By E. Derbyshire, K. J. Gregory and J.R. Hails. Pp.312. (Butterworths/Westview, Boulder, Colorado: 1979.) Hbk £12.50, $31.25; pbk $13.50.

It is now more than 30 years since the famous symposium, commemorating the centenary of the birth of W. M. Davis, marked the beginning of the so-called "new" geomorphology of which these two texts, in their differing ways, are examples. In 1952, in "The Dynamic Basis of Geomorphology" *(Bull. geol. Soc. Am.* **63**, 923–938) Strahler amplified the "dynamic approach" advocated at the symposium, which required the quantitative determination of landform characteristics and the formulation of empirical equations by the methods of mathematical statistics. This "counterbalance" which, he felt, should replace the Davisian explanatory descriptive method, involved

> the study of gravitational and molecular shear stresses acting on the materials of the earth's surface, which behave as elastic or plastic solids, or viscous fluids to produce the characteristic varieties of strain or failure that constitute weathering, erosion, transportation or deposition.

Process in Geomorphology and *Geomorphological Processes* both adopt this approach but with differing degrees of emphasis and with differing success. The first is almost a caricature where the landform characteristics appear frequently to be forgotten and mathematical statistics and empirical equations seem, at times, to be incorporated for their own sakes rather than to make an understanding of the operation of processes easier. Indeed, some of the authors have succeeded in making relatively simple topics seem difficult. Even final-year undergraduates will have difficulty with parts of this volume, especially if they have no background in physics and hydraulics. The excellent introductory chapters go some way towards helping such students. *Geomorphological Processes* succeeds in adopting a rigorous approach but in a more balanced way, maintaining a clarity of mathematical as well as of verbal expression which illuminates rather than confuses.

In *Process in Geomorphology* two other defects mar what is, nonetheless, a useful compendium of scholarly essays. The idea that landforms may encapsulate the results of many processes, of which those now seen in operation are but the most recent, finds no place in the book. Changes of climate and of process over time, for example those resulting from the advances and retreats of Pleistocene ice sheets, are largely ignored. Second, rates of operation of processes, particularly in the context of the rare event, are inadequately considered by most of the authors. The concluding chapter by Dr Thornes does nothing to remedy these defects.

Geomorphological Processes similarly concentrates on contemporary processes and to some degree suffers from the same defects. However, the short introductory chapter makes clear that this is a conscious choice and indicates the need to examine these other facets of landform developments, at greater length, elsewhere.

Derbyshire, Gregory and Hails succeed in making their study an exercise in of the landscape, by constantly maintaining the link between process and form. This is made possible by the arrangement of the book. Essentially, processes are divided into glacial and periglacial, fluvial and marine with only a small chapter devoted to the work of the wind. A much shorter introduction than in the Embleton and Thornes volume allows each of the three major sections to be treated at greater length. The effect is most marked in Hails's chapter which is a full and authoritative treatment of coastal processes. By contrast, Clark in *Process in Geomorphology,* asked to attempt a similar geomorphology, rather than in the physics exercise in half the space, not unnaturally makes a more superficial presentation.

These two volumes are interesting for contrasts in style rather than in approach. Both have value. *Process in Geomorphology* will be useful to the readers to whom it is directed much more as a reference work than as a course primer; *Geomorphological Processes* could well satisfy the opposite function. □

C. Kidson is Professor of Physical Geography and Head of the Department of Geography at The University College of Wales, Aberystwyth.

Sedimentary, my dear Watson!

Bruce Sellwood

Origin of Sedimentary Rocks. 2nd Edn. By H. Blatt, G. Middleton and R. Murray. Pp.782. (Prentice-Hall: 1980.) £20.10, $30.95.

UNTIL the middle of this century the study of sedimentary rocks fell largely under the two headings of stratigraphy and petrology. Sedimentology as a separate integrative discipline emerged no more than 30 years ago and a major surge in the subject has occurred over the past 20 years, much of the impetus for new developments coming from the demands of the oil and mineral extraction industries.

One of the first integrative modern sedimentological texts was *Origin of Sedimentary Rocks* (Prentice-Hall, 1972) which stressed the genesis of sediments and the processes of physical and chemical sedimentation. At that time the book was generally welcomed as a well organized, comprehensive and admirably clear text. After eight years, a totally revised edition has appeared. The new edition is almost 25% longer than the original but retains the successful structure and stylistic clarity of its predecessor. The book is divided into six parts and contains 19 chapters, each of which ends with a useful annotated reference list.

Part I introduces the scope of sedimentology, emphasizing the role of field investigation. Discussion of various facets of the geological cycle is undertaken in Part II.

Part III is a review of terrigenous clastic sediments, and commences with a new and welcome chapter on the geochemistry of natural waters. This is followed by coverage of weathering processes and products, mineralogy of sandstones, cementation and sandstone diagenesis, and sandstone classification. The chapter on cementation and diagenesis is also new and reflects the recent developments in this field which stem directly from petroleum reservoir studies.

Part IV considers the origin, diagenesis and classification of limestones, dolomites and evaporites, while Part V tackles the disparate topics of chert, phosphates, iron-rich rocks and manganese nodules.

The final part is devoted to facies models. Although the nucleus of this section appeared in the first edition, it has been largely re-written and greatly expanded. It emphasizes the predictive nature of facies analysis and environmental interpretation and, although larger than the original version, is still briefer than it should be in view of the importance of accurate facies modelling to economic geology.

There are no attempts at facies modelling using electric-log analysis, which is a pity since these methods are so important in contemporary hydrocarbon exploration. Some diagrams in the first edition were unclear and have been re-drafted, but there are still a few over-reduced diagrams or diagrams taken from poor originals. Yet these are minor quibbles.

Overall, this is an excellent general source text for students and professionals. I hope it appears in paperback — soon.

Bruce Sellwood is a Lecturer in Geology at The University of Reading.

New mineralogy

D. Flinn

Principles of Mineral Behaviour. By A. Putnis and J. D. C. McConnell. Pp.257. (Blackwell Scientific/Elsevier/North-Holland: 1980.) Hbk £18, $45; pbk £9.80, $24.95.

THIS book is the first of a new Blackwell series entitled *Geoscience Texts;* it is also the first new mineralogy textbook for undergraduates to appear for a very long time. Each year a number of ostensibly new texts appear but on examination they turn out to be updated or merely rehashed versions of books which have been with us for decades, or even since mineralogy began. Descriptive mineralogies in one guise or another are legion, and optical crystallographies seem to appear every year, all closely modelled on *Optical Crystallography* by Whalstrom which first appeared in 1943 and now is in its fifth edition (Wiley, 1979). For years I have had to refer my students, for accounts of mineral behaviour, to odd chapters in texts written for physicists, ceramicists, metallurgists and lately geochemists.

Now, at last, such accounts have been put together in one book and brought bang up to date for geology undergraduates. The book is descriptive, being concerned with kinetics rather than equilibria, but the necessary minimum of thermodynamics is clearly and simply presented. The approach is process orientated rather than mineral orientated even though the processes are all demonstrated in common minerals. The authors concentrate on polymorphism and on the sub-solidus behaviour of solid solutions and show how the end products depend more on the time available than on equilibrium relations. I hope that in future publishers will use the resources they now devote to publishing new optical crystallographies to updating this type of textbook and expanding it to cover other types of mineral behaviour such as crystal growth and crystal deformation. □

D. Flinn is Professor of Geology at the University of Liverpool.

Extended senses

Robert N. Colwell

Remote Sensing: Optics and Optical Systems. By P.N. Slater. Pp.575. (Addison-Wesley: 1980.) $34.50, £20.70.

FOR years, remote-sensing educators and their students, as well as engineers and researchers in remote-sensing-related fields, have needed a comprehensive book which would explain, in a readily understandable way, both the principles and the design of various kinds of remote-sensing instruments, and the procedures best used in measuring the performance of such instruments. Professor Slater's book seeks to satisfy that need.

In this book, of 14 chapters and 9 appendices, a systematic treatment is given of the various optical and electro-optical systems that are currently used for remote sensing in the 0.4 to 16μm range, including the Landsat multispectral scanner, return beam vidicon and thematic mapper systems. (Microwave remote sensing — both active and passive — is reserved for another book in the series.) In particular, the author describes in detail the appropriate methods for correcting the "spectral signatures" of features for both atmospheric and system-dependent effects. A substantial amount of the author's own high-quality research is integrated into the book; however its greatest strength lies in the clarity with which Slater has explained, using little more than basic physics and simple mathematics, the relevant work of scientists in a great variety of fields, ranging from human optics to electromagnetic theory, from electronic circuitry to spectroradiometry, and from basic atmospherics to matter/energy interactions at the Earth's surface. Of particular value is his consistent use throughout the book of a powerful performance criterion, the "modulation transfer function".

Despite the prodigiousness of his undertaking, Slater is careful, for the most part, to cite references to the many principles with which his book deals. Thus it is all the more disconcerting when, on occasion, he makes flat assertions (without providing documentation) such as "the average eye can discriminate approximately 5 million colors but only 200 shades of gray". Actually there are almost as many (widely disparate) values quoted in the literature as there are experts who have addressed this highly important matter. An occasional citation is improperly listed in Slater's book but such errors will no doubt be corrected in future editions.

All told, Slater's book is highly authoritative, unusually up to date and comprehensive, clearly written, well fortified with relevant tables and neatly drawn figures, and eminently worth perusal by any serious remote-sensing scientist. ☐

Robert N. Colwell is a Professor at the University of California, Berkeley, and teaches courses in photo interpretation and remote sensing.

Beyond Wiener

Freeman Gilbert

Geophysical Signal Analysis. By E.A. Robinson and S. Treitel. Pp.466. (Prentice-Hall: 1980.) £23.40, $36.

GEOPHYSICAL signal analysis became a subject of research by the authors in the early 1950s, when they were graduate students at MIT. They were members of the Geophysical Analysis Group, sponsored by a consortium of petroleum and exploration companies, whose rather ill-defined goal was to use digital computers for the application of the ideas, techniques and methods of Norbert Wiener to the interpretation of reflection seismograms. In the intervening quarter-century the seismic exploration industry has "gone digital" in a big way. In a typical day the industry processes several thousand computer tapes (9 track, 1600 BPI, 2400 feet) of data on multiple mainframe computer systems. By contrast, in the early 1950s the authors used the Whirlwind II computer (now in the Smithsonian Museum). Input was on punched paper tape, output on electric typewriters and secondary storage on magnetic wire; instruction time was in milliseconds and memory was 1024 16-bit words. It was the premier computer of its day. The authors deserve a large part of the credit for developing many of the currently used computational procedures in geophysical signal analysis and their book provides the basis for understanding and using them. Their firsthand knowledge and experience are evident throughout.

The method of least squares pervades science and engineering; it is the central theme of the present volume. Basically the book is concerned with filtering, the process of removing or diminishing "noise" in order to reveal and enhance the "signal". With the reflection seismogram as a model, the authors develop least squares digital filtering from first principles in the first eight chapters. The subjects of minimum phase lag, minimum delay, causality and stability are presented in a thorough manner, and the concept of a prediction filter is introduced. The writing is well balanced between theory and application. This arises from the blending of the authors' talents — Robinson is a mathematician at home with geophysics and Treitel is a geophysicist at home with mathematics. They complement one another very well.

The second half of the book is devoted primarily to the subject of deconvolution, especially predictive deconvolution. The opening chapter is devoted to a discussion of stationary time series. The method of generalized harmonic analysis, attributed primarily to Kolmogorov and Wiener, is mentioned and briefly described, but the bulk of the chapter is devoted to Wold's predictive decomposition theorem, and deductions from it made by the authors. By using the innovational representation of a nondeterministic stationary time series (that is, the output of minimum delay filter whose input is a white noise process), the authors set the stage for the succeeding chapters on predictive deconvolution. Here, we learn how predictive deconvolution is used to remove multiple reflections and reverberations from seismic traces. The wave propagation theory is the familiar one-dimensional, plane wave theory. It does not conform to the usual common depth point stacking procedures and does not always give good results. The newer slant stacking procedures are better and give more nearly the plane wave stacks for which predictive deconvolution is designed.

The book ends with a chapter on spectral estimation. Both the ARMA (auto regressive, moving average) and ME (maximum entropy) methods are described.

Robinson and Treitel have produced a useful, carefully written book for students, teachers and prospectors. Many of the subjects that they consider are important beyond the realm of seismic prospecting for oil, and their book should find appeal among all scientists and engineers interested in digital signal processing. ☐

Freeman Gilbert is a Professor of Geophysics at the Institute of Geophysics and Planetary Physics, University of California, San Diego. In the early 1950s he was a graduate student at MIT and a member of the Geophysical Analysis Group.

GFD for all levels

D. Anderson

Geophysical Fluid Dynamics. By J. Pedlosky. Pp.624. (Springer-Verlag: 1979.) DM 83.50, $43.90. *Geophysical Fluid Dynamics for Oceanographers.* By J.J. von Schwind. Pp.307. (Prentice-Hall: 1980.) £19.45, $29.95.

THE science of physical oceanography is in the process of emerging from the dark ages, a process begun some 40 years ago and greatly accelerated in the past decade. The companion subject of dynamical meteorology has had a similar history but is slightly further advanced. Together they constitute geophysical fluid dynamics (GFD), an important new discipline which is attracting interest because of its relevance to improving weather forecasts and understanding climate.

With the expansion of the subject has come a demand for textbooks. The two reviewed here are intended for new graduate students or advanced undergraduates, but the levels of the courses for which the texts are suitable are totally different. To the novice student embarking on one of the rather compressed courses which tend to be offered at many institutions, Pedlosky would be rather daunting. Von Schwind, however, is sufficiently elementary that it could be picked up and read profitably by any new student. The subjects discussed in Chapters 3 and 5 are important — wind-driven ocean circulation, the explanation of western boundary currents and some aspects of large-scale wave dynamics — though the presentation could have been markedly improved if some indication of modern thinking had been included (only three of the references are post 1970). Much of the remainder of the book is in the nature of support material. Chapter 2, by far the largest (over half the length of the book), deals with various aspects of basic fluid mechanics and geophysical fluid mechanics. Since many students enter oceanography with no prior knowledge of fluids, this can be quite useful even if much of it is readily available elsewhere.

Pedlosky on the other hand is suitable for more advanced graduate students and specialists. Almost everyone in this category is bound to find something of interest in this intensive book each time he opens it, be it for casual reading or reference. The scope of the book could have been more catholic but the treatment of his chosen material is generally very thorough, one notable exception being equatorial dynamics, an important aspect of modern thinking which gets only marginal coverage in the last chapter (and none at all from von Schwind). The mid-latitude quasi-geostrophic potential vorticity equation is well explained and there is a good section on instability theory. The presentation is mathematically, rather than physically orientated but Pedlosky is a text one can recommend to all serious students of GFD.

D. Anderson is a Lecturer in Atmospheric Dynamics at Oxford University.

Unmathematical air

John Green

Atmospheric Physics. By J. V. Iribarne and H.-R. Cho. Pp.212. (Reidel: 1980.) Dfl. 40, $15.95.

ANOTHER text that is said to require only elementary mathematics — in this case to understand the terrestrial atmosphere. I must confess a bias. To my mind *no* book on physics should use more complex mathematics than is needed to make its point. Conversely, no text should go out of its way to avoid some essential mathematics because the reader may not understand it. He had better be made aware of his deficiency and of the intellectual benefit that will result from being able to understand the mathematical part of an argument.

The authors identify the need for a survey of the physics of the terrestrial atmosphere, suitable for students having completed only one year of a course in general physics. They give a general description of the structure of the atmosphere, atmospheric chemistry, radiation, thermodynamics, cloud physics, and atmospheric electricity and dynamics. Each topic is treated sympathetically but briefly. At the end of each chapter there are questions and problems designed to aid the learning process, and hints on solutions at the end of the book. Diagrams are mostly well drawn but the writing is occasionally careless.

The philosophy of presenting material without recourse to more than elementary mathematics is admirable. Execution is less so. The student is given the appropriate formula with only the most general guidance to the most comprehensive monographs where, I believe, he will find the derivation very difficult to follow. More detailed reference to a text of intermediate standard would help. I think it is useful to comment that a stated formula contains some of the variables you would expect and has a plausible functional form, besides giving a detailed reference to a succinct derivation in the same notation.

This quibble apart, our poor, unfortunate, non-mathematical person will find this a useful, interesting and balanced survey of atmospheric physics.

John Green is Reader in Physical Climatology and teaches a course in atmospheric physics at Imperial College of Science and Technology, University of London.

Star of the lecture theatre

Stephen P. Maran

Discovering the Universe. By Charles E. Long. Pp.511. (Harper & Row: 1980.) Pbk $16.50. *Dynamic Astronomy.* 3rd Edn. By Robert T. Dixon. Pp.523. (Prentice-Hall: 1980.) Pbk £11, $16.95. Teacher's manual also available.

AT A university in southern California, 500 undergraduates in a long, narrow auditorium observe a professor, dimly visible at centre stage, microphone around the neck, lecture on black holes, the Big Bang and life in space. At an East Coast university, a comparable throng files into two adjacent auditoria as the dividing wall between the rooms is electrically retracted; at the same time a large turntable at the front of one of the halls slowly rotates away the entomology instructor who has just finished a lecture. Into view spins the astronomer with demonstration table, microphone and piles of computer-graded examinations that are ready for return to the students. As the lecture begins and the lights fade away, a coloured photograph of a spiral galaxy appears overhead, projected in duplicate to provide identical side-by-side images for the benefit of those sitting at far left and far right in the wings of the assembly.

As the elementary, non-mathematical course in astronomy has become so popular in North America, perhaps 50 competing texts have appeared. The latest, *Discovering the Universe*, is alternately thoughtful and chatty, with many apropos anecdotes and numerous, seemingly-scrawled sketches that are as effective as the usual formal diagrams. At its best (taken on its own terms), *Discovering* introduces the unwary to general relativity with the aid of two inquisitive ants (Einstant and Newtant), or describes Mercury as seen by *Mariner 10*. In the latter case, only a few details are given, but much effort then follows in explaining what one might experience on a visit to the planet. The explanation is clear enough that the student will understand why each prospective phenomenon is anticipated. Marring the otherwise accurate text, however, are erroneous statements (each of which I have seen before) to the effect that *Uhuru* discovered Cygnus X-1 and that the optical pulsar in the Crab Nebula was discovered on the famous Lick Observatory rotating-shutter photographs. Those photos were obtained well after telegrams announced the discovery of the optical pulses at Steward Observatory, their confirmation at McDonald Observatory, and their identification with a specific star — the pulsar — at Kitt Peak National Observatory.

Dynamic Astronomy, here in a third edition, is distinguished by "flip pages", which bear in the margins sequences of diagrams that can be made to spring into life. Riffle them and you will observe Halley's Comet in retrograde motion, tail growing as it approaches the Sun, then shrinking as the comet hurtles back into the far reaches of the Solar System. Other sequences depict planetary, stellar and star cluster motions. Better yet, try turning the same pages slowly and you will find besides the diagrams a well-polished text, familiar by now to the students "in over two hundred colleges and universities". Yet improvements are still to be made. A section on "Observing Meteors" gives the wrong explanation for why more meteors are observed after midnight than before. Contrary to Fig.9-15, globular clusters are strongly concentrated to the centre of the Galaxy although they are more conspicuous in the halo. The sources of the gamma-ray bursts discovered by Klebesadel with data from the *Vela* satellites are confused with the very different "X-ray bursters" studied by Grindlay. Less significantly, although surprising in a mature text, the names of astronomers often are spelled wrong, including those of Olbers, Jastrow, Solomon and Maran. Somehow, when the reviewer's name is misspelled, it always is noticed. □

Stephen P. Maran is at the Laboratory for Astronomy and Solar Physics, NASA–Goddard Space Flight Center, Greenbelt, Maryland.

The Institute of Biology has published a *Handbook for the Animal Licence Holder*, edited by H.V. Wyatt, intended particularly for newcomers to laboratory animal work. The *Handbook* is available from 41 Queen's Gate, London SW75HU, price £2.50 or £2 for five or more copies.

Quantum mechanics for librarians

R. B. Jones

Simple Quantum Physics. By P. Landshoff and A. Metherell. Pp.177. (Cambridge University Press: 1980.) Hbk £12, $29.95; pbk £4.25, $9.95. *An Introduction to Quantum Physics.* By A. P. French and E. F. Taylor. Pp.670. (W.W. Norton/Nelson: 1979.) Hbk $17.95, £15.95; pbk $11.75, £8.95. *Principles of Quantum Mechanics.* By R. Shankar. Pp.612. (Plenum: 1980.) $29.50, £18.59.

AT THE end of any introductory quantum mechanics course, the lecturer sadly concludes that there is not yet a satisfactory textbook while the bemused undergraduate inscribes on the desk top "Schrodinger waives the rules — probably". The three books considered here do not solve the lecturer's problem although each has its merits. They are well produced with a few misprints and each is well provided with problems for which solutions are given at the end of the text. Apart from this they have little in common.

Simple Quantum Physics is based on lectures given at Cambridge University to mathematicians and physicists, and is suitable for a one-term introductory course. I enjoyed reading the book for its brevity and clarity. There is a very brief account of the basic ideas of wave mechanics in which the standard problems are treated up to the hydrogen atom without spin. The mathematical treatment is clear but so concise that only able students will follow what is going on. A short treatment of perturbation theory leads to the most distinctive feature of the book, an excellent introductory account of atomic absorption and stimulated emission and how these can be used to understand a laser. The final three chapters introduce solid-state quantum physics by way of band theory and the physics of real transistor devices. I will recommend this book to our college library and urge my students to read it; however I will use a more detailed text as the required reading in my own course of introductory quantum mechanics.

Out of the MIT Introductory Physics Series comes French and Taylor's *An Introduction to Quantum Physics.* Here the emphasis is on the experimental basis of quantum physics, the coverage ranging from the great historical experiments to modern laser spectroscopy. The basic approach is wave mechanical, although two chapters on photon polarization experiments point the way to the abstract notion of a state vector. All the basic problems up to the hydrogen atom are treated, and both spin and hydrogen fine structure are mentioned. Atomic radiation is discussed although no systematic perturbation theory is included. The mathematical treatment is unfortunately sparse. The student who uses this book will enjoy it, but he will have to do it all again. For that reason I cannot recommend it as a set text in British universities; however, in North America, where there is more time for an undergraduate degree, I would commend it highly for a one-term or one-semester course. Even in Britain many lecturers will want this book for its outstanding illustrations and examples, and libraries should have a copy for students.

Principles of Quantum Mechanics by Shankar is intended for a year-long first course in quantum mechanics. In contrast to the previous authors, Shankar introduces the subject in a loose axiomatic way using abstract linear spaces and the Dirac notation. The attempt to explain the subject at this level means that the book is very long, and even long-winded at times, without covering as much physics as one should expect in 600 pages or so. There are some compensations, however. There is a clear introduction to the Feynman path integral formalism and special attention is paid to the role of symmetries in quantum mechanics. All standard perturbation theory is covered, including fine structure effects in hydrogen. Angular momentum addition is treated and the concept of an irreducible tensor operator is introduced. Scattering theory, electromagnetic field quantization and even the Dirac equation are briefly discussed. This book has merits as a reference book, but I think its place is not in the lecture room. □

R. B. Jones is a Lecturer in Physics at Queen Mary College, University of London.

Approaches to nuclear reactions

James Lowe

Introduction to Nuclear Reactions. By G.R. Satchler. Pp.316. (Macmillan, London/Wiley, New York: 1980.) £23, $59.95. *The Physics of Nuclear Reactions.* By W.M. Gibson. Pp.338. (Pergamon: 1980.) Hbk £19.50, $45; pbk £6.95, $16.75.

BOTH of the authors of these two volumes set out to provide a textbook on nuclear reactions, as opposed to static nuclear properties, for the undergraduate student. Usually, a single book covering the whole of nuclear physics provides adequate detail for an undergraduate course, and several such books are available. There are, in fact, some advantages in having the entire subject in one book — the unity of the subject renders the division into static properties and reactions arbitrary and sometimes confusing. However, the extra detail covered in a specialist book may well be useful to the more inquisitive student, especially if he intends to pursue the study of nuclear physics beyond undergraduate level.

Two books apparently aiming to teach much the same general topic at the same level could hardly be more different in their content and approach. Satchler's book, after an excellent introductory chapter, and an extensive section on basic scattering theory, covers in detail all the theories and models which have formed the basis of our understanding of nuclear reactions over the past 20 years. The treatment is authoritative; many of these models have been developed by Satchler himself. However, the approach is by no means historically orientated — topics of current interest, such as heavy-ion scattering and multi-step processes, are discussed and the book contains plenty of examples of the application of the methods described to "real live" data. The descriptive material is excellent, although curiously the only point at which it is a little misleading is the discussion of the concept of cross-sections, which is sometimes difficult to get across to students. The scope of the book is broad, both in the depth of treatment and in the wide range of topics covered. For example, even Glauber theory, which is useful at energies much higher than those usually encountered in nuclear reaction studies, is described briefly.

By contrast, Gibson's book, which is a new edition of his book *Nuclear Reactions,* provides a much more concise account of the basics of nuclear reaction theories. Unfortunately, several topics which are an essential part of most undergraduate courses are omitted; for example, there is no quantitative treatment of direct reactions. The discussion is clear throughout, and Gibson seems sometimes to assume a more elementary starting point for his readers than does Satchler. Gibson provides an excellent discussion of particle accelerators (though not of any other aspect of techniques — a curious selection), and also gives an account of the physical principles of fission and fusion reactors. His final chapters on nuclear forces and nuclear spin give good accounts of these topics, with a breadth of treatment beyond that directly relevant to nuclear reactions.

The choice between these books for the student must be based on their content. The student who intends to concentrate on pure nuclear physics will find Satchler's treatment valuable, and the content of this book will be useful well into the student's postgraduate years. In fact, the treatment is sufficiently extensive and not too elementary for many research physicists. For the student who intends to move towards applied nuclear physics or energy research, especially if involved in the physics of reactors, Gibson's book may be the more relevant. □

James Lowe is Senior Lecturer in Nuclear Physics in the Department of Physics, University of Birmingham, UK.

Astrophysics

R.J. Tayler

The Physics of the Interstellar Medium. By J.E. Dyson and D.A. Williams. Pp.194. (Manchester University Press/Halsted: 1980.) £11.95, $24.95. *Radiative Processes in Astrophysics.* By G.B. Rybicki and A.P. Lightman. Pp.382. (Wiley-Interscience: 1979.) £16.95, $39.85.

ASTRONOMY, as well as being an exciting subject in its own right, gives a student an excellent opportunity to see how different branches of physics are brought together to solve a particular problem. For that reason, stellar structure, which involves atomic and nuclear physics, thermodynamics and statistical physics, electromagnetism, fluid dynamics and gravitation, is a popular option for physics undergraduates. The physics of the interstellar medium provides another excellent example of the interplay of different branches of physics and chemistry, such as the origin and transfer of radiation throughout the electromagnetic spectrum, hydrodynamics and magnetohydrodynamics and the physics and chemistry of small particles and molecules. Although it is a good subject for an undergraduate course, there is a lack of suitable textbooks. There is one outstanding book, *Physical Processes in the Interstellar Medium* by L. Spitzer (Wiley-Interscience, 1978), but this is far too advanced for the average or weak undergraduate.

The book by Dyson and Williams is written for the undergraduate audience and contains a range of topics which almost exactly matches the lecture course given to third-year undergraduates at Sussex. To some extent there is emphasis on the research interests of the two authors, interstellar hydrodynamics (Dyson) and interstellar grains and the formation and destruction of molecules (Williams), but overall it gives a good impression of the range of physical problems met in the interstellar medium. It is clearly written and contains a useful number of examples, and, as it uses SI units, it should be immediately intelligible, at least to UK undergraduates. It is a pleasure to be able to recommend to my students a book at a level suitable for all of them.

Rybicki and Lightman are also concerned with astronomy as physics but instead of dealing with a particular branch of astronomy they have considered all astrophysical applications of one branch of physics, the origin and transfer of radiation which is of particular importance in the interstellar medium. This is an outstanding book which has reached me when I am preparing an MSc course on the subject. It is making my task both simpler and more enjoyable, although I regret the need to translate its formulae from cgs to SI units.

Rybicki and Lightman have covered a wide range of topics, including the fundamentals of radiative transfer and a review of both electromagnetism and special relativity as well as discussing the various types of radiation from free particles (bremsstrahlung, synchrotron and inverse Compton radiation) and atomic and molecular structure and spectral lines. They have succeeded admirably in giving all the necessary information without too much unnecessary detail and the book contains many problems with worked solutions. The book should be of wide general interest to physicists as well as astronomers, since only the problems are specifically astronomical, and I hope that its title will not prevent them from recognizing its value. □

R.J. Tayler is Professor of Astronomy at the University of Sussex, managing editor of Monthly Notices of the Royal Astronomical Society *and author of three introductory astronomical textbooks.*

Pure Che[illegible] Library
SCIENCE COLLEGE
92 Upper Circular Road
CALCUTTA

Facets of thermal physics

P. T. Landsberg

Principles of Thermodynamics. By J. A. Beattie and I. Oppenheim. Pp.329. (Elsevier Scientific: 1979.) Dfl. 98, $47.75. *Thermal Physics.* 2nd Edn. By C. Kittel and H. Kroemer. Pp.473. (W.H. Freeman: 1980.) $22.95, £11.30. Instructor's manual also available.

THE images of joint authors BO and KK (for short) transmitted by these two interesting and worthwhile books are rather different. BO (viewed as a unit) dresses in a formal dark suit, is reliable, a trifle old-fashioned and, because of a personal communication of 1911 acknowledged in the text (p.170), not a hot-headed youngster by any means. He has a good sense of the history of the subject: nineteenth century references are frequent, and only one or two literature citations from the 1970s occur — "new-fangled stuff compared with Clausius and Joule" one hears him say. He likes open systems and chemical potentials: van't Hoff's equilibrium box is not often discussed. Indeed there is an interest in fundamentals: Carathéodory's principle, the unattainability principle, Gibbs paradox are all considered though the commentaries and developments of these topics in the 1970s are not touched on. "Too new" said the author, replacing his pince-nez.

KK, on the other hand, wears a sports jacket and jeans, is most excited by newer developments, and readily departs from general principles to discuss effects and applications in this (largely rewritten) second edition: superconductors, solar constant, Overhauser effect, p-n junctions, heat pumps and particle–antiparticle equilibrium are all noted. Some of these interesting things are, however, in the problems. "Come on boys, this is fascinating stuff, you can do it", he says, while writing his concluding piece on negative temperatures. But there must always be a doubt on that point. An author may not even be able to solve his own problems if he has left the subject for a while.

If our authors should meet, one can imagine BO asking KK "Why is Nernst not mentioned or unattainability or Pfaffians?". "Well, there are special books for that" answers KK, puffing at his pipe; "Why don't you write one like that?". "I have" says BO.

Knowing some of the authors, the appearances inferred from their books are actually quite wrong. But the point is clear: thermal physics is a many-faceted subject and to get the full flavour it may not be enough to read one book. Even two, the above two, lucky though one is to have them, may not be enough. One would have to go elsewhere to learn about negative heat capacities, Saha's equation or black-hole thermodynamics.

This is not the place to make criticisms of details, but it may be of wider interest to note the following point (taken from my own book on the subject), which neither author stresses though they both deal with the zeroth law of thermodynamics. This law is $(A \equiv B, A \equiv C) \Rightarrow (B \equiv C)$. It states the transitivity of thermal equilibrium (denoted by $\equiv$), A, B and C being systems. Now comes a question which in its simplicity is suitable for the promotion board of the Scientific Civil Service: does the existence of a single temperature function follow, such that temperature equality and thermal equilibrium is the same? The answer is "No". This may be seen by interpreting "$\equiv$" as the condition for parallelism of lines, A, B, C in space. This requires *two* angles which specify the directions of the lines to be equal. Hence we have shown that cases exist in which a single function is inadequate. □

P. T. Landsberg is Professor of Mathematics at the University of Southampton.

Introductory . . .

W. Cochran

Elements of Solid State Physics. By M. N. Rudden and J. Wilson. Pp.186. (Wiley: 1980.) Hbk £14.50, $40; pbk £5.95, $17.

CHARLES Kittel maintains that "One test of a solid state course is that it should treat holes correctly". The book under review fails this test, since the authors cheerfully discuss the Fermi level in terms of the water level in a flask, a hole in terms of a bubble in the water, and that is almost their last word on the subject. However this is not meant to be a book for professional physicists, who will (perhaps) know that the wave vector of a hole state is *minus* the wave vector of the corresponding unoccupied electron state, a fact that cannot be of much interest to an engineer designing a large-scale integrated circuit. The authors are not lacking in ambition, for they have set out to write a textbook for the student who requires to be introduced to the Bohr model of the hydrogen atom on p.15, and who is yet presumed to be ready to understand the physics of the tunnel diode by p.153 — they are writing for

> the increasing number of students studying solid state physics at an introductory level, many of whom require only sufficient background to understand the operation of modern solid state electronic devices.

I began to read the book with some scepticism but I became increasingly impressed by the ability of the authors to convey ideas by means of arguments and analogies which, although they do not treat each concept "correctly", are almost always helpful and will not have to be consciously unlearnt by the student who goes on to study the subject more deeply. Once or twice I thought the authors were unhelpfully wrong, as when they write on p.87 that "for an intrinsic semiconductor at absolute zero the Fermi level coincides with the top of the valence band", then continue on p.88 with an unconvincing reason for its being "halfway between the two bands", and finally on the same page give a valid argument why it "can perhaps be best understood" to occupy the latter position.

All in all, the authors are to be congratulated on their success in writing a non-mathematical book on solid-state physics which is genuinely an introduction. Possibly students of electrical engineering need not know more than this about solid-state physics *per se* but, if they must, they will find several other books in the reading list, including inevitably Charles Kittel's *Introduction to Solid State Physics* (5th Edn, Wiley, 1979), which is close to the other end of the spectrum of introductions. □

W. Cochran is Professor of Natural Philosophy at the University of Edinburgh.

. . . and applied solid-state physics

J. C. Anderson

Introduction to Applied Solid State Physics: Topics in the Applications of Semiconductors, Superconductors, and the Nonlinear Optical Properties of Solids. By R. Dalven. Pp.330. (Plenum: 1980.) $27.50, £17.33.

THE philosophy of this book is that solid-state physics teaching, in physics courses, often lacks reality — a view with which I would not disagree.

The book is based on a lecture course, given by the author in the Department of Physics at the University of California, Berkeley. The course evidently set out to *survey* the *physics* involved (author's italics) in a variety of solid-state devices. To this end, the first two chapters rapidly review basic semiconductor physics and the p–n junction, giving liberal references to *Introduction to Solid State Physics* (5th Edn, Wiley, 1979) by Charles Kittel (who apparently encouraged the author to develop his course) and other relevant texts. The following chapters then deal with semiconductor devices such as the transistor (both bipolar and field effect), metal–semiconductor junctions and their applications such as Schottky diodes, and the present generation of devices available through MOS technology. A chapter devoted to other semiconductor devices is something of a rag-bag and rather unwisely attempts to encapsulate the physics of amorphous semiconductors in four pages. In this fast-moving field such an attempt is decidedly unwise. However the day is always saved at the end of each chapter by a suggested reading list and references to key papers, which should enable the reader to overcome any shortcomings of the exposition. The remaining three chapters deal with generators and detectors of electromagnetic radiation, superconductive devices and materials and, finally, nonlinear optical properties of solids.

Essentially, the book is offering itself to lecturers wishing to illuminate their solid-state physics teaching, and to those students who might have sufficient curiosity about the relevance of their solid-state physics course to go and read a book. In this context the book succeeds, as evidently did the author's lecture course. In a wider context it is also useful as a quick reference guide to a scientist or engineer wishing to remind himself, or learn for the first time, of the basic physics behind a device with which he has to deal. It successfully dispels, *inter alia*, the delusions that a squid is a kind of cuttle fish or that the Josephson junction is a crossroads near Bethlehem. □

J. C. Anderson is Professor of Electrical Materials in the Department of Electrical Engineering, Imperial College of Science and Technology, University of London.

Electromagnetism, old and new

John Pain

Electricity and Magnetism. 3rd Edn. By W.J. Duffin. Pp.467. (McGraw-Hill: 1980.) Pbk £7.50, $18. *Fundamentals of the Theory of Electricity.* By I.E. Tamm. Pp.684. (Mir/Central Books, London: 1980.) £6.95.

FORTY years separate the first editions of these two books yet the essential contents are common to both. Most topics in physics advanced beyond all recognition during this particular period but the foundations of electricity and magnetism, as laid by the scientists of the nineteenth century, remain definitive.

Tamm is one of the great names of Russian physics; here an engaging slip in the biographical blurb awards him the Noble (*sic*) Prize. His book was first published in 1925 and ran to eight editions with only minor amendments before he died in 1971. His final preface acknowledged the need for a major revision but the ninth edition, published after his death, left the work unchanged. The book under review is a translation of that posthumous edition and the preserved style of the original anchors it firmly in the period of the inter-war years. This will commend it to a diminishing group of readers — those with a sense of history and a nostalgia for the European teaching tradition of that time. A lucid translation and good Romanian printing may encourage a wider audience but the theoretical bias, unrelieved by reference to practice, plus the use of an unfashionable system of units, will make no appeal to the physics undergraduate who has grown up in the world of radio, television, computers and communication satellites.

When Tamm first wrote this book radio was still in its infancy and he was under no pressure to include those experimental details which illustrate the extraordinary power of electromagnetic theory. An unrevised version of the same book is difficult to justify in 1981.

If Tamm's is the work of a past master, Duffin's is a model of what a modern textbook should be. The third edition deserves to be even more successful than the first two, for the author has expanded some of the more advanced topics and revised his presentation, helped by imaginative typography. This really is a book for all physics students (and course lecturers). The basic theory is there but solid physics flesh covers the mathematical bones without running to the fat of an over-written text. The book covers the first two years of an honours course in a British university and Duffin's technique is to take the early chapters slowly and to accelerate when the student is well grounded. His concern not to lose the reader is shown by commentaries and by revision summaries at the end of each chapter which bring the preceding pages into sharper focus. In addition there are nearly 300 problems and the answers are given to all of them with many useful hints.

There are few books which leave the impression that an outstanding lecture course has been transferred from blackboard to page — this is one of them. ☐

John Pain is Reader in Physics at the Imperial College of Science and Technology, University of London.

Learning soil physics

David L. Rimmer

Soil Physics. By T. J. Marshall and J. W. Holmes. Pp.345. (Cambridge University Press: 1980.) Hbk £33, $77.50; pbk £10.95, $21.95. *Applied Soil Physics: Soil Water and Temperature Applications.* By R. J. Hanks and G. L. Ashcroft. Pp.159. (Springer-Verlag: 1980.) DM 39.50, $22.20.

FOR some time an alternative has been required to the standard text on soil physics, *Soil Physics* by Baver *et al.* (4th Edn, Wiley, 1972), especially as it never appeared in paperback and eventually became too expensive for undergraduates to buy. The book by Marshall and Holmes is thus to be welcomed. In it they present a full account of the subject, concentrating in the first half on fundamental and theoretical aspects, especially of soil water, and in the later chapters dealing with the more field-orientated or management aspects, for example structural stability, compaction, tillage and irrigation.

The material is readable and attractively presented; tables and figures break up the pages, avoiding the problem of page after page of continuous text interspersed only by intimidating equations. Throughout there are many references; about 600 in total, these are brought together at the back of the book to provide a most useful reference list. There are, however, certain inconsistencies between chapters. For example, in Chapter 2 the total potential of soil water is defined in terms of pressure, osmotic and gravitational potentials only, where pressure potential is taken to cover matric potential under certain conditions. In Chapter 12 the total potential of water in plants is defined in terms of all four potentials separately. The choice of symbols and the use of SI units is generally good. Units for density however cause some problems; these appear in different places as kg m^{-3}, g cm^{-3} and Mg m^{-3}. Yet these are relatively minor criticisms of what is a very good book, which should become widely adopted.

Applied Soil Physics by Hanks and Ashcroft is different in many respects, although aimed at a similar readership (students of soil physics and scientists in related fields). First, its scope is more limited, as suggested by its sub-title *Soil Water and Temperature Applications.* Second, it specifically aims to show how the basic theoretical equations can be applied to field situations by means of numerous worked examples. These examples are backed up by sets of problems at the end of each unit (chapter). The aspects of soil water are dealt with in four units and soil temperature in a final unit. By working through the units the physical significance of the rather abstract theoretical relationships is made more apparent. The book is not strictly a textbook, but then it was not intended to be such (as evidenced by references to relevant textbooks at the end of each unit). What the book offers is an aid to teaching the subject. Its usefulness as such is somewhat diminished by the use of non-SI units, which may cause problems for students. But as a source of worked examples and relevant problems it will be of value to teachers of soil physics. ☐

David L. Rimmer is a Lecturer in the Department of Soil Science at the University of Newcastle upon Tyne.

Pedology illustrated

Milo I. Harpstead

Soils: Their Formation, Classification and Distribution. By E.A. FitzPatrick. Pp.353. (Longman: 1980.) £25, $60.

SOIL science has attracted increased public awareness in recent years and has become the professional concern of scientists in a wide variety of fields. Soils can be lost by erosion and contaminated by pollutants, sometimes with disastrous consequences, and vary in their suitability to produce crops, support structures and accept waste products. Soils are being inventoried on a world-wide scale and soil survey reports are now consulted for purposes of land-use planning.

This book covers a wide range of topics

text books
from
MACMILLAN
PRESS

ORBITALS & SYMMETRY
D. S. URCH
This book introduces both symmetry and molecular orbital theory in one volume. The mathematics which chemists need to apply to symmetry ideas is developed using molecular examples, (rather than abstract mathematical ones). At all stages worked problems are given in the text and further exercises are set at the end of each chapter.
The aim has been to give the reader a critical appreciation of two very important aspects of theoretical chemistry, molecular orbital theory and symmetry, and to show him how to apply them to a wide range of chemical problems.
£4.95, 0 333 26675 7

CHEMISTRY FOR THE LIFE SCIENCES
Second Edition
J. C. DAWBER and
A. T. MOORE
The book is primarily designed as an integrated chemistry text for undergraduate students on degree courses in biology, biochemistry, pharmacy, and medicine at universities, polytechnics, and colleges of higher education.
It will also serve as a text for HNC and HND students in Applied Biology and Medical Laboratory subjects, and students studying for MIBiol. and Membership of the Institute of Medical Laboratory Technicians. Students on joint honours courses (e.g., Combined Science) in which biology is a major subject will find the book useful.
£16.00, 0 333 25820 7
£8.95, 0 333 25821 5

PLANT METABOLISM
H. D. KUMAR and
H. N. SINGH
The highly complicated, interconnected metabolic cycles that continue unceasingly within the living cells of plants are refreshingly simplified in this classic text for students of botany and those studying agricultural botany and plant physiology in agricultural colleges.
The physiologist, the biochemist, and the phytochemist – each will find his particular field ably covered in this masterful analysis which integrates the different features of plant chemicals – their potential functions, their synthesis and degradation – into a comprehensive account.
£4.95, 0 333 25638 7

Macmillan Press books are available from leading booksellers
If you have difficulty obtaining these titles please write to Tony Watson, Macmillan Retail Ltd, Houndmills, Basingstoke, Hants, enclosing purchase price + 70p (p&p).

M MACMILLAN PRESS

normally taught in undergraduate courses. It is well illustrated with excellent photographs and drawings; of particular interest are the illustrations of landscape–soil relationships for several climatic regions of the world and the numerous photomicrographs from thin-sections showing the soil micromorphology resulting from the processes being described. In addition, there are 16 colour plates of soil profiles, representing major soil regions of the world, and a number of tables which represent an exhaustive compilation of information from international sources.

However, the book does have its faults. Mr FitzPatrick has reviewed the world literature on soils, but surprisingly he seems to have missed the point in his critical assessment of Hans Jenny's opinion about the factors of soil formation. Less importantly, he cites secondary sources for the theories of Davis and Penck on slope development, giving the impression that they were formulated several decades later than they were.

Many authors have developed their own nomenclature similar to that of Mr FitzPatrick; such terminology is usually short-lived and confusing if it does not reflect broad acceptance within the discipline. The author reminds us that soils form a continuum in space and time and that "this situation defies classification". He contends that for 80 years pedologists have tried unsuccessfully to construct a satisfactory hierarchical system of soil classification and this leads him to conclude that it simply cannot be done. The USDA uses the most elaborate hierarchical system in the world and so receives Mr FitzPatrick's severest criticism, an assessment which is contentious, to put it mildly.

This is a valuable book for pedologists, but has limited value for students who are attempting to master an established system of soil nomenclature and classification. □

Milo I. Harpstead is Professor of Soil Science at the University of Wisconsin, Stevens Point, and the author of Soil Science Simplified *(Iowa State University Press, 1980.)*

All round food production

J. C. Bowman

Agricultural Ecology: An Analysis of World Food Production Systems. By G. W. Cox and M. D. Atkins. Pp.721. (W. H. Freeman: 1979.) £15.70, $29.95.

A RIGOROUS analysis of world food production systems involves an understanding of many aspects of science and social science. One method of teaching agriculture is to consider the relevant disciplines of chemistry, biology, economics and so on, and then gradually to bring this knowledge together to show how it relates to food production systems adopted on farms. Undoubtedly this approach tests the patience of students and requires them to have faith that those aspects of science and social science which they study are relevant and will eventually interrelate to give a sound understanding of agriculture. Students can gain confidence and make a more purposeful approach to the choice of constituent subjects if they have an overview of agriculture from the start of their studies.

Cox and Atkins have made a brave and largely successful attempt to provide such an overview by considering agriculture as a group of ecosystems. Their book is in three parts. The first deals with the ecological and historical context of agriculture, the second with the dynamics of agroecosystems and the third with agriculture and the future. For a text with such a wide range of subjects which require integration and focus there is a big advantage in the easy and consistent style which the authors have achieved. There is a great deal of quantitative material which has been made comprehensible, even enjoyable.

It is perhaps inevitable that with only two authors some imbalance of subject matter has arisen. In particular there is considerable emphasis on agricultural pest problems — nearly half of Part Two is devoted to this topic — and there is much more attention given to crop rather than livestock production systems. This may represent, albeit inadvertently, the authors' view of the future importance of these subjects. A stronger element of the anthropological and sociological aspects of agriculture throughout the world, particularly in appreciating changes in agriculture in the developing world, would have been beneficial. Even so, the text gives a good account of the ecological approach to agriculture.

The book is profusely illustrated with photographs, diagrams and tables. Some of the diagrams have a minimum of explanation and in several cases students will need to return to the source for a full appreciation of what they portray. Each chapter has a good bibliography which draws heavily on American literature, and contains relatively little reference to publications after 1976 even though the book was not published until 1979.

The book should be valuable to all students of food production systems throughout their course. A text to be read and re-read as a student gains knowledge of the constituent parts of the subject. □

J. C. Bowman is Director of the Centre for Agricultural Strategy at the University of Reading.

Conserving resources

Ronald F. Hensler

Soil and Water Conservation (For Productivity and Environmental Protection). By F.R. Troeh, J.A. Hobbs and R.L. Donahue. Pp.718. (Prentice-Hall: 1980.) £16.85, $25.95.

THE first two chapters of this readable and well-edited book give an historical as well as a world-wide perspective on soil erosion. As well as providing students with a broad view of their subject, they should be required reading for those making decisions which in any way determine how land will be used, and thus to what degree that land will be subject to erosion by water and wind.

With the principles laid down, the authors go on to evaluate the parameters involved in water and wind erosion forces and soil-loss predictions. References here, as throughout the volume, are a useful combination of the key original concepts, recent research, and review articles.

Chapters 8, 9, 10 and 12 cover soil and plant management and conservation structure techniques (tillage, cropping systems, terraces, dams and so on) that can be employed to reduce erosion. Critical design requirements as well as discussions of the pros and cons of each technique are provided. The remainder of the book deals with soil surveys and land-use planning, mine and construction site reclamation, pasture, range and forest management, water conservation, drainage, irrigation, soil and water pollution, economics, United States soil and water conservation agencies, and conservation around the world.

Other features include a summary, (though not of the thought-provoking variety), and references at the end of each chapter. The illustrations are appropriate, and the subheadings have been well planned with each subsection able to stand by itself so that a teacher can assign portions of chapters without losing the integrity of the whole.

The major difference between this and other books on soil and water conservation is that most textbooks are engineering-orientated to meet the on-the-job needs of a practising soil conservationist. This book, then, is particularly welcome in filling a gap.

Soil and Water Conservation should certainly be considered for adoption by teachers of soil and water conservation courses. In addition, it will be a useful reference book for all agronomists, all politicians and all who make decisions that affect land use. I only hope such people read it. □

Ronald F. Hensler is Associate Professor of Soil Science at the University of Wisconsin, Stevens Point.

BOOKS RECEIVED

Mathematics

BELLMAN, R. Analytic Number Theory: An Introduction. Mathematics Lecture Note Series, No.57. Pp.195. ISBN 0-8053-0360-X. (Benjamin/Cummings, Reading, Massachusetts: 1980.) $19.50.

FREEDMAN, H.I. Deterministic Mathematical Models in Population Ecology. Pp.254. ISBN 0-8247-6653-9. (Marcel Dekker: 1980.) SwFr. 64.

JONES, W.B. and THRON, W.J. Continued Fractions. Analytic Theory and Applications. Encyclopedia of Mathematics and its Applications, Vol.11. Pp.428. ISBN 0-201-13510-8. (Addison-Wesley: 1980.) $37.50.

KRISHNAIAH, P.R. (ed.). Developments in Statistics, Vol.3. Pp.254. ISBN 0-12-426603-7. (Academic: 1980.) $35.

ROBINSON, S.M. (ed.) Analysis and Computation of Fixed Points. Proceedings of a Symposium conducted by the Mathematics Research Centre, The University of Wisconsin, Madison, May 1979. Pp.413. ISBN 0-12-590240-9. (Academic: 1980.) $22.50.

TREVES, F. Introduction to Pseudodifferential and Fourier Integral Operators. Vol.2, Fourier Integral Operators. Pp.649. ISBN 0-306-40404-4. (Plenum: 1980.) Np.

Astronomy

ANDREW, B.H. (ed.). Interstellar Molecules. IAU Symposium No.87 held at Mont Tremblant, Québec, Canada, August 1979. Pp.706. Hbk ISBN 90-277-1160-7; pbk ISBN 90-277-1161-5. (Reidel: 1980.) Hbk Dfl. 145, $76.50; pbk Dfl. 65, $34.

DE JAGER, C. The Brightest Stars. Geophysics and Astrophysics Monographs. An International Series of Fundamental Textbooks, Vol.19. Pp.458. Hbk ISBN 90-277-1109-7; pbk ISBN 90-277-1110-0. (Reidel: 1980.) Hbk Dfl. 140, $73.50; pbk Dfl. 60, $31.50.

DRYER, M. and TANDBERG-HANSSEN, E. (eds). Solar and Interplanetary Dynamics. Symposium No.91 held in Cambridge, Massachusetts, August 1979. Pp.558. Hbk ISBN 90-277-1162-3; pbk ISBN 90-277-1163-1. (Reidel: 1980.) Hbk Dfl.125, $66; pbk Dfl.55, $28.95.

HALLIDAY, I. and McINTOSH, B.A. (eds). Solid Particles in the Solar System. Symposium No.90, organized by the IAU and held at Ottawa, August 1979. Pp.442. Hbk ISBN 90-277-1164-X; pbk ISBN 90-277-1165-8. (Reidel: 1980.) Hbk Dfl.95, $49.95; pbk Dfl.50, $26.50.

Physics

FLEAGLE, R.G. and BUSINGER, J.A. An Introduction to Atmospherics. 2nd Edn. International Geophysics Series, Vol.25. Pp.432. ISBN 0-12-260355-9. (Academic: 1980.) $29.50.

FRÜNGEL, F.B.A. Sparks and Laser Pulses. Vol.IV, High Speed Pulse Technology. Pp.488. ISBN 0-12-269004-4. (Academic: 1980.) $49.50.

HASS, G. and FRANCOMBE, M.H. (eds). Physics of Thin Films, Vol.11. Pp.336. ISBN 0-12-533011-1. (Academic: 1980.) $42.

LIOU, K.-N. An Introduction to Atmospheric Radiation. Pp.392. ISBN 0-12-451450-2. (Academic: 1980.) $32.50.

LUCOVSKY, G., PANTELIDES, S.T. and GALEENER, F.L. (eds). The Physics of MOS Insulators. Proceedings of the International Topical Conference, North Carolina State University, Raleigh, June 1980. Pp.369. ISBN 0-08-025969-3. (Pergamon: 1980.) Np.

SHANNON, R.R. and WYANT, J.C. (eds). Applied Optics and Optical Engineering, Vol.VIII. Pp.407. ISBN 0-12-408608-X. (Academic: 1980.) $46.

TREDER, H.-J., VON BORZESZKOWSKI, H.-H., VAN DER MERWE, A. and YOURGRAU, W. Fundamental Principles of General Relativity Theories: Local and Global Aspects of Gravitation and Cosmology. Pp.216. ISBN 0-306-40405-2. (Plenum: 1980.) Np.

WALLS, D.F. and HARVEY, J.D. (eds). Laser Physics. Proceedings of the 2nd New Zealand Summer School. Pp.287. ISBN 0-12-733280-4. (Academic: 1980.) $32.50.

Chemistry

EBERLE, A. and WIELAND, T. (eds). Perspectives in Peptide Chemistry. Pp.444. ISBN 3-8055-1297-X. (Karger, Basel: 1981.) SwFr. 145, DM 174.

FREISER, H. (ed.). Ion-Selective Electrodes in Analytical Chemistry, Vol.2. Modern Analytical Chemistry Series. Pp.291. ISBN 0-306-40500-8. (Plenum: 1980.) Np.

MORRISON, S.R. Electrochemistry at Semiconductor and Oxidized Metal Electrodes. Pp.401. ISBN 0-306-40524-5. (Plenum: 1980.) Np.

SCHOFIELD, K. Aromatic Nitration. Pp.376. ISBN 0-521-22362-3. (Cambridge University Press: 1981.) £27.50.

SMEDLEY, S.I. The Interpretation of Ionic Conductivity in Liquids. Pp.195. ISBN 0-306-40529-6. (Plenum: 1980.) Np.

THOMAS, J.M. and LAMBERT, R.M. (eds). Characterisation of Catalysts. Pp.283. ISBN 0-471-27874-2. (Wiley-Interscience: 1981.) £16.50, $49.

Technology

CRIPPS, Y.M. Controlling Technology: Genetic Engineering and the Law. Pp.154. ISBN 0-03-056806-4. (Preger, New York: 1980.) Np.

FELDMAN, S.L. and WIRTSHAFTER, R.M. On the Economics of Solar Energy: The Public–Utility Interface. Pp.256. ISBN 0-669-03449-5. (Lexington Books, Farnborough, Hants: 1981.) £16.50.

GRAHAM, D. and THOMSON, J. Grenz Rays: An Illustrated Guide to the Theory and Practical Application of Soft X-rays. Pp.149. ISBN 0-08-025525-6. (Plenum: 1980.) £10.

KEENEY, R.L. Siting Energy Facilities. Pp.413. ISBN 0-12-403080-7. (Academic: 1980.) $32.

STRUB, A.S. and EHRINGER, H. New Ways to Save Energy. Commission of the European Communities. Proceedings of the International Seminar held in Brussels, October 1979. Pp.1252. ISBN 90-277-1078-3. (Reidel: 1980.) Dfl.135, $71.05.

STRUB, A.S. and IMARISIO, G. Hydrogen as an Energy Vector. Commission of the European Communities. Proceedings of the International Seminar held in Brussels, February 1980. Pp.703. ISBN 90-277-1124-0. (Reidel: 1980.) Dfl.95, $50.

STRUB, A.S. and UNGEMACH, P. (eds). Advances in European Geothermal Research. Proceedings of the 2nd International Seminar held in Strasbourg, March 1980. Pp.1086. ISBN 90-277-1138-0. (Reidel: 1980.) Dfl.120, $63.

YIH, C.-S. (ed.). Advances in Applied Mechanics, Vol.20. Pp.235. ISBN 0-12-002020-3. (Academic: 1980.) $31.50.

Biological Sciences

CLAPHAM, A.R. (ed.). The IBP Survey of Conservation Sites: An Experimental Study. International Biological Programme 24. Pp.344. ISBN 0-521-22697-X. (Cambridge University Press: 1980.) £25.

CUNHA, T.J. Horse Feeding and Nutrition. Pp.292. ISBN 0-12-196560-0. (Academic: 1980.) $26.50.

GALE, R.P. and FOX, C.F. (eds). Biology of Bone Marrow Transplantation. Proceedings of the 1980 ICN–UCLA Symposia, Keystone, Colorado, February 1980. Pp.566. ISBN 0-12-273960-4. (Academic: 1980.) $40.

GRESSER, I. (ed.) Interferon 1980, Vol.2. Pp.99. Flexi ISBN 0-12-302251-7. (Academic: 1980.) £6.80, $16.

GUTHRIE, D.M. Neuroethology: An Introduction. Pp.221. Flexi ISBN 0-632-00303-0. (Blackwell Scientific: 1980.) £8.75.

HAITH, M.M. Rules That Babies Look By: The Organization of Newborn Visual Activity. P.146. ISBN 0-85859-033-7. (Lawrence Erlbaum Associates, Hillside, New Jersey: 1980.) $16.50.

HALTENORTH, T. and DILLER, H. A Field Guide to the Mammals of Africa Including Madagascar. Pp.400. ISBN 0-00-219778-2. (Collins: 1980.) £8.95.

HANSEN, H.H. and RØRTH, M. Lung Cancer 1980. Postgraduate Course. International Congress Series No.525. II, World Congress on Lung Cancer, Copenhagen, June 1980. Pp.152. ISBN 90-219-0447-0. (Excerpta Medica: 1980.) $39, Dfl.80.

HARRIS, H. The Principles of Human Biochemical Genetics. 3rd revised edn. Hbk ISBN 0-444-80264-9; pbk ISBN 0-444-80256-8. (Elsevier/North-Holland Biomedical: 1980.) Hbk $65, Dfl.135; pbk $26.75, Dfl.55.

HENDRY, B. Membrane Physiology and Cell Excitation. Pp.159. Hbk ISBN 0-7099-0148-8; pbk ISBN 0-7099-0149-6. (Croom Helm: 1981.) Hbk £9.95; pbk £4.50.

HUNT, R.K. (ed.). Neural Development. Part 1, Emergence of Specificity in Neural Histogenesis. Pp.427. ISBN 0-12-153115-5. (Academic: 1980.) $34.

IETSWART, J.H. A Taxonomic Revision of the Genus Origanum (Labiatae). Leiden Botanical Series, Vol.4. Pp.153. Flexi ISBN 90-6021-464-3. (Kluwer: 1980.) Dfl.60.

ISHII, S., HIRANO, T. and WADA, M. (eds). Hormones, Adaptation and Evolution. Proceedings of the International Symposium in Honour of Professor Hideshi Kobayashi, Tokyo, July 1979. Pp.343. ISBN 3-540-10033-4. (Springer-Verlag: 1980.) DM69, $40.80.

JOHNSON, M.H. (ed.). Development in Mammals, Vol.4. Pp.327. ISBN 0-444-80274-6. (Elsevier/North-Holland: 1980.) $49.75, Dfl.102.

KOGAN, M. and HERZOG, D.C. (eds). Sampling Methods in Soybean Entomology. Pp.587. ISBN 3-540-90446-8. (Springer-Verlag: 1980.) DM88, $52.

KREIER, J.P. (ed.). Malaria. Vol.1, Epidemiology, Chemotherapy, Morphology, and Metabolism. Pp.416. ISBN 0-12-426101-9. (Academic: 1980.) $49.

LARCHER, W. Physiological Plant Ecology. 2nd Totally Revised edition. Pp.303. Flexi ISBN 3-540-09795-3. (Springer-Verlag: 1980.) DM59, $34.90.

LEAF, A. and COTRAN, R.S. Renal Pathophysiology. 2nd Edn. Pp.410. Hbk ISBN 0-19-502688-8; pbk ISBN 0-19-502689-6. (Oxford University Press: 1980.) Hbk np; pbk £7.95.

MEADOWS, R. Pocket Atlas of Human Histology. Pp.459. Hbk ISBN 0-19-261248-4; pbk ISBN 0-19-261177-1. (Oxford University Press: 1980.) Hbk £15; pbk £7.50.

MIFLIN, B.J. (ed.). The Biochemistry of Plants: A Comprehensive Treatise, Vol.5. Amino Acids and Derivatives. Pp.669. ISBN 0-12-675405-5. (Academic: 1980.) $65.

MILLER, T.A. (ed.). Neurohormonal Techniques in Insects. Pp.282. ISBN 3-540-90451-4. (Springer-Verlag: 1980.) DM79, $46.70.

MORRIS, R.H., ABBOTT, D.P. and HADERLIE, E.C. Intertidal Invertebrates of California. Pp.690. ISBN 0-8047-1045-7. (Stanford University Press: 1980.) $30.

PIAGET, J. Adaptation and Intelligence. Organic Selection and Phenocopy. Pp.124. ISBN 2-7056-5852-1. (University of Chicago Press, London: 1980.) £6.60.

PIZZEY, G.A. A Field Guide to the Birds of Australia. Pp.460. ISBN 0-00-219201-2. (Collins: 1980.) £12.50.

POON, L.W. *et al.* (eds). New Directions on Memory and Aging. Proceedings of the George A. Talland Memorial Conference. Pp.572. ISBN 0-89859-035-3. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $36.

PYCOCK, C.J. and TABERNER, P.V. (eds). Central Neurotransmitter Turnover. Pp.197. ISBN 0-7099-0471-1. (Croom Helm: 1980.) £11.95.

RANDALL, L.L. and PHILIPSON, L. (eds). Virus Receptors Pt1. Bacterial Viruses. Receptors and Recognition Series B, Vol.7. Pp.147. ISBN 0-412-15660-1. (Chapman & Hall: 1980.) £15.

ROBERTS, C. Science, Animals and Evolution: Reflections on Some Unrealized Potentials of Biology and Medicine. Contributions in Philosophy, Number 14. Pp.221. ISBN 0-313-21479-4. (Greenwood, London: 1980.) £10.95.

ROY, A.K. and CLARK, J.H. Gene Regulation by Steroid Hormones. Pp. 316. ISBN 3-540-90464-6. (Springer-Verlag: 1980.) DM84, $49.60.

SCHULZ, I. *et al.* (eds). Hydrogen Ion Transport in Epithelia. Developments in Bioenergetics and Biomembranes, Vol.4. Proceedings of the International Conference held in Frankfurt, July 1980. Pp.474. ISBN 0-444-80290-8. (Elsevier/North-Holland Biomedical: 1980.) $70.75, Dfl. 145.

SCOTT, W.A. *et al.* (eds). Mobilization and Reassembly of Genetic Information. Miami Winter Symposia, Vol.17. Pp.459. ISBN 0-12-633360-2. (Academic: 1980.) $35.

SCRIABINE, A., LEFER, A.M. and KUEHL, F.A. Jr (eds). Prostaglandins in Cardiovascular and Renal Function. Monographs of the Physiological Society of Philadelphia, Vol.6. Proceedings of the 20th Annual A.N. Richards Symposium of the Physiological Society of Philadelphia, Pennsylvania, May 1978. Pp.498. ISBN 0-85200-532-6. (MTP, Lancaster: 1980.) £29.95.

SLEIGHT, P. (ed.). Arterial Baroreceptors and Hypertension. Pp.540. ISBN 0-19-261259-X. (Oxford University Press: 1980.) Np.

STUMPF, P.K. (ed.). The Biochemistry of Plants: A Comprehensive Treatise, Vol.4. Lipids: Structure and Function. Pp.693. ISBN 0-12-675404-7. (Academic: 1980.) $65.

VAN DER STARRE, H. (ed.). ECRO—ISOT VII. Joint Meeting held at "de Leeuwenhorst" Congress Centre, Noordwijkerhout, the Netherlands, July 1980. Proceedings of the International Symposium on Olfaction and Taste and of the 4th Congress of the European Chemoreception Research Organization. Pp.50. ISBN 0-904-14720-7. (IRL, London & Washington: 1980.) £22, $50.

VINCENT, J.F.V. and CURREY, J.D. (eds). The Mechanical Properties of Biological Materials. Symposia of the Society for Experimental Biology, No. XXXIV. Pp.513. ISBN 0-521-23478-6. (Cambridge University Press: 1980.) £30.

von BITTER, P.H. and MERRILL, G.K. Naked Species of *Gondolella* (Conodontophorida): Their Distribution, Taxonomy, and Evolutionary Significance. Life Sciences Contributions, Royal Ontario Museum, Number 125. Pp.50. Pbk ISBN 0-888-266-6. (Royal Ontario Museum, Toronto: 1980.) $3.

WEICK, F. and BROWN, L.H. Birds of Prey of the World. Pp.159. ISBN 0-00-219277-2. (Collins: 1980.) £15.

WELLS, R.M.G. Invertebrate Respiration. The Institute of Biology's Studies in Biology No.127. Pp.71. Flexi ISBN 0-7131-2806-2. (Edward Arnold: 1980.) £2.25.

WILLIAMS, J.G. A Field Guide to the Birds of East Africa. Pp.415. ISBN 0-00219179-2. (Collins: 1980.) £7.95.

Applied Biological Sciences

ABLIN, R. J. (ed.) Handbook of Cryosurgery. Pp.424. ISBN 0-8247-6981-3. (Dekker: 1980.) SwFr.128.

ALBERTINI, A. (ed.) Radioimmunoassay of Hormones, Proteins and Enzymes. International Congress Series No.528. Proceedings of the International Symposium,Gardone Riviera, May 1980. Pp.272. ISBN 90-219-0454-3. (Excerpta Medica: 1980.) $56, Dfl.115.

BAUMANN, N. (ed.). Neurological Mutations Affecting Myelination. Research Tools in Neurobiology, Correlations to Human Neurological Diseases. Inserm Symposium No.14. Proceedings of the International Symposium held in Seillac, France, April 1980. Pp.565. ISBN 0-444-80270-3. (Elsevier/North-Holland: 1980.) $85.75, Dfl.176.

CARTENSEN, J. T. Solid Pharmaceutics: Mechanical Properties and Rate Phenomena. Pp.259. ISBN 0-12-161150-7. (Academic: 1980.) $35.

CAWLEY, J. C., BURNS, G. F. and HAYHOE, F. G. J. Hairy-Cell Leukaemia. Recent Results in Cancer Research, Vol.72. Pp.123. ISBN 3-540-09920-4. (Springer-Verlag: 1980.) DM 56, $33.10.

CIBA FOUNDATION SYMPOSIUM 76 (new series). Environmental Chemicals, Enzyme Function and Human Disease. Pp.380. ISBN 90-219-4082-5. (Excerpta Medica: 1980.) $63.50, Dfl.130.

CLEMMESEN, J. *et al.* (eds.) Quantitative Aspects of Risk Assessment in Chemical Carcinogenesis. Archives of Toxicology, Supplement 3. Symposium held in Rome, April 1979. Pp.336. Flexi ISBN 3-540-09584-5. (Springer-Verlag: 1980.) DM 58, $34.30.

COHEN, P. Recently Discovered Systems of Enzyme Regulation by Reversible Phosphorylation: Molecular Aspects of Cellular Regulation, Vol.1. Pp.273. ISBN 0-444-80226-8. (Elsevier/North-Holland: 1980.) $63.50, Dfl.130.

COHN, W. E. (ed.). Progress in Nucleic Acid Research and Molecular Biology, Vol.24. Pp.279. ISBN 0-12-540024-1. (Academic: 1980.) $32.50.

COSTA, M. Metal Carcinogenesis Testing: Principles and In Vitro Methods. Pp.164. ISBN 0-89603-017-2. (Human Press, Clifton, New Jersey: 1980.) $29.50 ($39.50 outside North America).

DI CHIARA, G. and GESSA, G. L. (eds). Glutamate as a Neurotransmitter. Advances in Biochemical Psychopharmacology, Vol.27. Pp.452. ISBN 0-89004-420-1. (Raven: 1980.) $41.

DUNCAN, W. (ed.). Thyroid Cancer. Recent Results in Cancer Research, Vol.73. Pp.142. ISBN 3-540-09328-1. (Springer-Verlag: 1980.) DM 60, $35.30.

EWERT, J. -P. Neuroethology: An Introduction to the Neurophysiological Fundamentals of Behavior. Pp.342. ISBN 3-540-09790. (Springer-Verlag: 1980.) DM 49, $29.

FIECHTER, A. (ed). Advances in Biochemical Engineering, Vol.16. Pp.148. ISBN 3-540-09807-0. (Springer-Verlag: 1980.) DM 72, $42.50.

FUMAGALLI, R., KRITCHEVSKY, D. and PAOLETTI, R. (eds). Drugs Affecting Lipid Metabolism. Symposia of the Giovanni Lorenzini Foundation, Vol.7. Proceedings of the VIIth International Symposium, Milan, May 1980. Pp.396. ISBN 0-444-80283-5. (Elsevier/North-Holland Biomedical: 1980.) $63.50, Dfl.130.

GARATTINI, S. *et al.* (eds). Advances in Pharmacology and Chemotherapy, Vol.17. Pp.326. ISBN 0-12-032917-4. (Academic: 1980.) $36.

GOLDBERG, E.P. and NAKAJIMA, A. (eds). Biomedical Polymers: Polymeric Materials and Pharmaceuticals for Biomedical Use. Proceedings from a Symposium held in April 1979 at the First Joint Congress of the American Chemical Society and the Chemical Society of Japan in Honolulu, Hawaii. Pp.457. ISBN 0-12-287580-X. (Academic: 1980.) $32.

GUSTAFSSON, J.-Å et al. (eds). Biochemistry, Biophysics and Regulation of Cytochrome P-450. Developments in Biochemistry Vol.13. Proceedings of the Third European Meeting held in Saltsjöbaden, Sweden, June 1980. Pp.626. ISBN 0-444-80282-7. (Elsevier/North-Holland: 1980.) $79.50, Dfl.163.

HAYAISHI, O., ISHIMURA, Y. and KIDO, R. (eds). Biochemical and Medical Aspects of Tryptophan Metabolism. Developments in Biochemistry, Vol.16. Proceedings of the 3rd International Meeting of the International Study Group for Tryptophan Research, Kyoto, Japan, August 1980. Pp.373. ISBN 0-444-80297-5. (Elsevier/North-Holland Biomedical: 1980.) $58.50, Dfl.120.

HOCUTT, C. H. *et al.* (eds). Power Plants: Effects on Fish and Shellfish Behavior. Pp.346. ISBN 0-12-350950-5. (Academic: 1980.) $25.

JACKOBY, W. B. (ed.). Enzymatic Basis of Detoxication, Vols I and II. Vol.I, pp.415, ISBN 0-12-380001-3; Vol.II, pp.369, ISBN 0-12-38002-1. (Academic: 1980.) Vol.I $43; Vol.II $28.50.

JAKOBY, W. B. (ed.). Enzymatic Basis of Detoxication. Vol.II, Biochemical Pharmacology and Toxicology Series of Monographs. Pp.369. ISBN 0-12-380002-1. (Academic: 1980.) $38.50.

KOEBNER, H. K. Lasers in Medicine, Vol.1. Pp.274. ISBN 0-471-27602-2. (Wiley-Interscience: 1980.) £39.50, $125.

KÖSTER, H. (ed.) Nucleic Acids Synthesis: Applications to Molecular Biology and Genetic Engineering. Nucleic Acids Symposium Series No.7. Proceedings of the International Symposium held in Egestorf, GFR, May 1980. Pp.396. ISBN 0-904-147-27-4. (Information Retrieval, Oxford: 1980.) £17, $40.

KROON, A.M. and SACCONE, C. (eds). The Organization and Expression of the Mitochondrial Genome. Developments in Genetics, Vol.2. Proceedings of the 12th International Bari Conference held in Martina Franca, Italy, June 1980. Pp.451. ISBN 0-444-80276-2. (Elsevier/North-Holland: 1980.) $69.75, Dfl.143.

LEVI-MONTALCINI, R. Nerve Cells, Transmitters and Behaviour. Proceedings of a Study Week at the Pontifical Academy of Sciences, October 1978. Pp.679. ISBN 0-444-80243-6. (Elsevier/North-Holland: 1980.) $109.75, Dfl.225.

LEVITT, J. Responses of Plants to Environmental Stesses. 2nd Edn. Vol.II, Water, Radiation, Salt, and Other Stresses. Pp.607. ISBN 0-12-445502-6. (Academic: 1980.) $45.

LUCARELLI, G., FLIEDNER, T. M. and GALE, R. P. (eds). Fetal Liver Transplantation: Current Concepts and Future Directions. Proceedings of the 1st International Symposium, Pesaro, Italy, September 1979. Pp.325. ISBN 90-219-0462-4. (Excerpta Medica: 1980.) $61, Dfl.125.

McLAREN, D.S. Nutritional Ophthalmology. 2nd Edn. of Malnutrition and the Eye. Nutrition: Basic and Applied Science. Pp.431. ISBN 0-12-484240-2. (Academic: 1980.) £32 (UK only), $74.50.

Psychology

NICKERSON, R.S. (ed.). Attention and Performance VIII. Proceedings of the 8th International Symposium, Princeton, New Jersey, August 20–25 1978. Pp.822. ISBN 0-89859-038-8. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $49.95.

PAPERT, S. Mindstorms: Children, Computers, and Powerful Ideas. Pp.230. ISBN 0-85527-163-9. (Harvester: 1981.) £9.95.

RIEBER, R. and SALZINGER, K. (eds). Psychology: Theoretical–Historical Perspectives. Pp.364. ISBN 0-12-588265-3. (Academic: 1980.) $21.

SEGAL, B., HUBA., G.J. and SINGER, J.L. Drugs, Daydreaming, and Personality: A Study of College Youth. Pp.254. ISBN 0-89859-042-6. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $19.95.

SNOW, R.E., FEDERICA, P.-A. and MONTAGUE, W.E. (eds). Aptitude, Learning, and Instruction. Vol.1, Cognitive Process Analyses of Aptitude. Vol.2, Cognitive Process Analyses of Learning and Problem Solving. Vol.1, pp.354, ISBN 0-89859-043-4; Vol.2, pp.336, ISBN 0-89859-046-9. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $49.95 (2 vol. set).

THOMPSON, R.F., HICKS, L.H. and SHVYRKOV, V.B. (eds). Neural Mechanisms of Goal-Directed Behavior and Learning. Pp.639. ISBN 0-12-688980-5. (Academic: 1980.) $49.50.

History of Science

ALBRITTON, C.C. Jr The Abyss of Time: Changing Conceptions of the Earth's Antiquity after the Sixteenth Century. Pp.249. ISBN 0-87735-341-7. (Freeman, Cooper: 1980.) $12.75.

nature

26 February 1981

15/2/82

Darwin's death in South Kensington

What is the British Museum (Natural History) for? Research? Public enlightenment, even entertainment? Or should it be, as Dylan Thomas said in another connection, a "museum that should be in a museum"? In spite of the stream of correspondence in the past several months, itself a measure of the museum's claim on public esteem and even affection, we are none the wiser. The museum, now at the beginning of its centenary year, has not replied to the several issues which have been raised since Dr Colin Patterson's cool comment on the role of cladistics in taxonomy (*Nature* **288**, 430; 1980). Halstead (this issue, page 742) is right to draw attention to this deafening silence.

The issues are not, after all, so complicated. Cladistics is a technique used in systematics for the classification of organisms that may or may not belong to distinct species. Cladism, of which Halstead is a well known enemy, is presumably the belief that cladistic analysis yields meaningful results. Halstead's original complaint against the museum was that a new public exhibition at the museum on "fossil man" is shot through with cladism to such a degree that familiar relationships between fossil human forms are obscured, which only went to show that the museum had been captured by Marxist ideology. The second charge, always a red herring, cannot be sustained. The public face of the museum is trendy, not revolutionary. The questions which have been raised, and which the museum must in the public interest answer, are not nearly as grand but are, by the same test, not so easily ignored.

Even the row there has been about cladism in public exhibitions is probably only a symbol for a more mundane but more important discontent — the museum's exhibition policy that has been evolving for the past five years or so. Waterhouse's Victorian pile has been a hive of activity throughout that period. The dinosaurs and stuffed birds have taken a back seat, or have been hidden away in back rooms, to make way for much more jazzy exhibitions. Although some of these are second-rate (that on ecology), simple inspection shows that others (that on human biology, subtitled "An exhibition of ourselves") fascinate and delight the crowds of visitors, many of them young, who are forever pressing buttons to provoke yet another physiological model into action. Nobody disputes the public demand for this kind of exhibition and there is even a possibility that it may assist in the education of the young and the general public. The practical issue is whether the museum should be encouraged, or even allowed, to fill itself with "exhibits" designed to "present the whole of modern biology" when the acknowledged consequence will be the replacement of "almost all" of the existing exhibitions.

This quarrel has been simmering for a long time, but the museum's dilemma is not negligible. Its postwar managers inherited a labyrinth of poorly lit galleries in which specimens of various kinds were so crammed together that their identity could hardly be discovered, especially if the ink had faded on the labels. The choice was to make the existing exhibition intelligible or to embark on a more ambitious policy. In 1972, when the pressures on public funds were magically almost imperceptible, the die was cast and the decision taken to remodel the whole museum. Since then, the museum has been constantly in hot water over its building plans. (The latest scheme for rebuilding the back of the museum, in the process changing the outline and the roof profile, confirms the suspicion that only public authorities can now commit vandalism openly.) Two questions persist. It there is a public need in London of a popular museum of science, why should its creation mean the end of the more scholarly exhibition created by the Victorians? The truth is that there is a need for both. And if the museum is irretrievably launched on its new policy of public education, how sensible is the course it is following?

The museum's showing so far is not encouraging. The latest brochure from the museum starts from the view that learning should be enjoyable. There is nothing wrong with that, even though the repetition of the jingle "learn and enjoy" may make many serious teachers wince. The danger, which many popular museums of science have discovered, is that enjoyment often interferes with learning. Pressing buttons to make things happen is an absorbing occupation at most ages, but is not by itself an educative process. In reality, the new exhibitions at the Natural History Museum, novel though they are for central London, differ only in scale from others to be found elsewhere in the world. It would be easier to believe that the museum is embarked on a worthwhile course if there were signs that it had seriously set out to break new ground in the public exhibition of science. Is it too late to look for signs of redemption in this respect? The new exhibitions promised for the summer may provide some clues.

The charge that the museum's new exhibitions are shot through with heresy, overstated though it may have been, also needs to be answered more categorically than in the past few weeks. Again there is right and wrong on each side. In the latest brochure, the popular euphemism for cladism is the phrase "the groups-within-groups classification", presumably a synonym for species. There is also the following passage: *Biologists try to reconstruct the course of evolution from the characteristics of living animals and plants and from fossils, which give a time scale to the story. If the theory of evolution is true* . . . If the words are to be taken seriously, the rot at the museum has gone further than Halstead ever thought. Can it be that the managers of the museum which is the nearest thing to a citadel of Darwinism have lost their nerve, not to mention their good sense? Or is it that somebody has calculated that the museum will increase its annual intake of visitors by enticing in scoffing creationists?

This passage nicely illustrates what has gone wrong at the museum. The new exhibition policy, the museum's chief interaction with the outside world, is being developed in some degree of isolation from the museum's staff of distinguished biologists, most of whom would rather lose their right hands than begin a sentence with the phrase "If the theory of evolution is true . . ." Nobody disputes that, in the public presentation of science, it is proper whenever appropriate to say that disputed matters are in doubt. But is the theory of evolution still an open question among serious biologists? And, if not, what purpose except general confusion can be served by these weasel words?

The most urgent need at the museum is to restore the once close relationship between the museum's scholarship and its external face. It is demeaning that many of those working as scientists at this distinguished place should not have been persuaded to harness their scholarship to the brave new policy. No doubt, if they had been more closely involved, the pace of development would have been slower and the museum might have been dissuaded from its promise to put "the whole of modern biology" on display. That, if the managers think about it, would be an advantage. For all its distinction, the museum is not equipped for such a task. What it can do well, as in the past century, is to tell the absorbing tale of natural history. It should bend its mind to that, not to worrying about cladisim.It should beware of selling out on Darwinism. And it should say, from time to time, what it is about. Hitherto, the museum has been too secretive.

0028-0836/81/090735-01$01.00

New Pentagon rules on overseas students

Secrecy threat to sensitive research plans

Washington

Restrictions on the use of foreign research students on non-classified research projects sponsored by the Department of Defense are in the offing. This prospect stems from the requirement, recently emphasized by Congress, that the multi-million dollar research programme sponsored by the department in very high speed integrated circuits should be strictly controlled to prevent the export of technical data of potential military value.

Pentagon officials say they are "bending over backwards" to minimize interference with the freedom of university research in applying controls. These controls were introduced in 1972 as the International Traffic in Arms Regulations (ITAR) to cover the export of unclassified technical data. Many academics, however, see a danger of renewed threats to academic freedom previously raised by secret military research on university campuses.

Because of the uncertainties which abound, the Defense Department has issued guidelines to both universities and private contractors carrying out sponsored research on integrated circuits. Where research can be classified as "process development" and is likely to lead to a specific improvement of military technology, controls will apply and it will be necessary to obtain an export licence from the State Department before the research is even discussed with a foreign citizen. But where the research can be classified as the "general pursuit of knowledge", with no particular application in mind, then controls will not apply.

Pentagon officials admit that there is a "grey area" dividing the two, particularly in materials science research and other areas closely linked to improving the efficiency of semiconductor systems. The Defense Department has told universities of its "preference" that foreign students should not be employed on integrated circuit programmes, but that the programme office will "make a decision based on the nature of the research" where restraint is not feasible.

The consequences could be significant. The proportion of foreign research students in US universities has risen dramatically in the past few years to about 50 per cent — and is particularly high in fields such as computer science, where first degree graduates can command large salaries in the private sector.

Several universities are concerned that strict interpretation of the regulations could reduce flexibility in research programmes, especially since the Department of Defense is the major supporter of this line of research. There is also mounting anxiety that the regulations might be abused. They are written so broadly that it is a federal crime to discuss with a foreign scientist any research result which might improve the "state of the art" of US military technology without prior approval by the State Department.

These fears have been reinforced by the National Security Agency's attempts to restrict the dissemination of cryptography research results obtained by Dr George Davida of the Georgia Institute of Technology (see *Nature* 19 February, p.621). Integrated circuit research is less sensitive, but government officials could theoretically use the regulations to close down semiconductor research in universities, on the grounds that, almost by definition, such research must advance the state of the art.

Commercial contractors for the Defense Department are also worried, since a strict interpretation of the regulations could prevent them from communicating research data, even if unclassified, to overseas affiliates.

The Defense Department, well aware of such objections and keen to avoid fresh campus confrontations such as those during the anti-war demonstrations of the 1960s, hopes that its new guidelines will convince universities of its concern for the freedom of research. Its proposals have been circulated through the Association of American Universities, a group of top US research universities, and will be discussed at a meeting of university and government officials in Washington next week.

Faced with other financial problems, universities are keen not to rock the boat by rejecting the new requirements as unreasonable. But many are wary of what they are letting themselves in for — hoping that it is not the thin end of an increasingly sticky wedge. **David Dickson**

British gene company opens books

Britain's national biotechnology company, Celltech, has begun marketing its first product: the monoclonal antibody to leukocyte interferon which, fixed to beads in a fractionation column, can purify interferon 5,000-fold in a single step. But since the discovery (see *Nature* **285**, 446-458; 1980) fell before November 1980, when Celltech negotiated rights to Medical Research Council biotechnical inventions, the National Research Development Corporation will retain a slice of the action. Profits on sales will be divided between Celltech, the corporation and the council.

The researchers who created the cell line, Dr David Secher of the MRC Laboratory of Molecular Biology, Cambridge, and Professor Derek Burke of the University of Warwick, also stand to gain but by different routes. In Burke's case, the standard revenue-sharing agreement will give the University of Warwick half the development corporation's profit; and Burke and colleagues in his laboratory will receive half of that, less the university's costs in running the laboratory for the relevant period. Secher, on the other hand, must look to his employer, the Medical Research Council. The council's agreement with Celltech allows for some reward for the researcher, but it is not clear at this stage how much. There is a standard arrangement for "inventors" in which a joint council-corporation panel meets every three years (the next meeting is in December 1982) to distribute largesse to those who have contributed to the council's coffers by work "outside normal duties". (In 1979, £130,000 was distributed to 30 inventors in amounts ranging from £40 to some thousands of pounds.) The catch is that many commercially valuable inventions — like monoclonal anti-interferon — fall within normal duties, and so in principle outside the inventors' scheme. Nevertheless, the council says it will interpret the phrase liberally in future.

In the end, the central issue is not how the loot will be divided but whether there will be any loot at all. The council hopes that the fledgling and hungry Celltech will be more assiduous than its traditional partner, the National Research Development Corporation, in exploiting inventions. Proposals are already in hand with a number of potential customers interested in the purification and radioimmunoassay of interferon using the antibody. However, Hoffmann la Roche is also believed to have developed similar cell lines, so there may be commercial competition.

Then the question will be how good is the Celltech line? There is more than one kind of interferon, and Celltech's antibody is strictly an antibody to a single interferon, one from the group of interferons supplied by the Wellcome Laboratories (consisting of 8 to 16 proteins) which formed Burke and Secher's starting material. Nevertheless, says Burke, evidence is accumulating that their antibody will bind to similar sites in a number of related interferons, so its specificity may not be unduly narrow. The initial quantities of antibody will be produced in Cambridge. Ultimately production will move to Celltech's new laboratories in Slough. Celltech expects a modest market for interferon for clinical trials over the next few years, with the promise of a much larger market if interferon proves to have therapeutic value.

Robert Walgate

0028-0836/81/090736-01$01.00

Celltech's deal with MRC

The Medical Research Council revealed last week the broad outlines of its five-year agreement with Celltech, signed on 6 November last. The agreement raises perturbing prospects of delays in the disclosure of scientific information. It concerns work done by researchers directly employed by the council and not that of its grant-holders employed by universities.

The council has undertaken to inform Celltech of commercially interesting discoveries by its researchers in fields related to biotechnology before the work is published. (In other fields, the traditional relationship with the National Research Development Corporation will be maintained.) The council says that the freedom of its research workers to communicate with other academic scientists will not normally be impaired, but occasionally publication would have to be delayed "for a short period" while a decision is taken on patenting.

Dr Sydney Brenner, director of the Laboratory of Molecular Biology and a member of Celltech's Science Council, is unconcerned. Delays will not be serious, and the issues involved are less difficult than those which arise when individuals have links with commercial companies.

In return for the right to information (and ultimate exploitation), Celltech will give the council a "substantial proportion" of any royalties or profits that ensue. The council will place this money in a separate "Celltech fund". When the fund reaches a sensible level, a subcommittee will call for proposals from both universities and council establishments on how it should be spent. In the early stages, when the fund is limited, grants will be restricted to fundamental research in areas related to biotechnology. If the money rolls in, the scope of the fund will be widened.

If, however, the sum in the fund rises beyond a fixed limit, which has been agreed between the Department of Education and Science and the council, but which neither will reveal because of its quasi-commercial nature, the dispensation that allows the council to accept money from Celltech will have to be renegotiated. Hitherto, the arrangement between the research councils and the development corporation has been that the research councils can receive no such payment. **Robert Walgate**

Polish universities

Sit-in complete

Last week, Poland's month-old student sit-in in Lodz ended with the signing of a wide-ranging agreement on academic autonomy. The students' action had begun over alleged delays by the Ministry of Science, Higher Education and Technology in registering the new Independent Students' Association (NZS). Later, as the Lodz University students were joined by those from the Lodz Technical University and Medical School, and then by students in Poznan, Krakow and Warsaw, their package of "postulates" grew to more than fifty issues.

The problem of getting NZS registered had been dragging on since last November, when a limited sit-in at Warsaw University won a promise that the ministry would produce the necessary regulations by 20 December. Difficulties with the statutes of NZS were finally resolved last week, by the inclusion of a clause stating that NZS would abide by the constitution of the Polish People's Republic, but without specific mention of the "leading role" of the Party.

The major part of the Lodz Accord, however, sums up progress on the promised liberalization of Polish academic life. Several clauses therefore simply reiterate what has already been conceded. These include: the abolition of "reserved" places at universities, granted at the discretion of rectors or ministry officials to candidates whose performance in the entrance examinations did not merit their admission; university courses to be decided by the staff of the universities concerned; staff appointed on academic merits only; student participation in university government, and the election of university rectors by the academic senate in secret ballot.

Of the new demands raised at Lodz, the most radical have been dropped, including the reduction of military service and the abolition of compulsory courses on Marxism. (Some choice in "sociopolitical subjects" has, however, been conceded.) One major gain for the students concerned language courses. This had been strongly opposed by the minister, Dr Janusz Gorski, who had argued that a university is not geared to teaching foreign languages *ab initio*. Nevertheless, there will now be a wider choice of languages to be studied as subsidiary subject, in effect ending the compulsory Russian which many science students in particular resented. Having had several years of Russian at school, they say they would like to study a Western language, to have greater access to current research literature.

Access to information should also be facilitated by clauses promising larger editions of textbooks (most students now have to make do with secondhand texts), more foreign currency for buying foreign journals and research apparatus and more open availability of existing literature.

A new system of student grants "more in tune with the principles of social justice" is promised for 15 May, when it will be submitted to the students for consultation. By May, too, the detested "practical training" in manual labour will be replaced by paid voluntary physical work.

Student participation in university senates and faculty councils is set at 20–33 per cent, the student delegates to have full voting rights except in the awarding of academic degrees and titles.

The students also seem to have won their demand that the police should not be able to enter university premises at will. This clause, however, will presumably have to undergo some modification before the new legislation on higher education finally comes before the Sejm (Parliament), if campuses are not to provide an unintended sanctuary for, say, a purse-snatcher on the run. Inclusion of the clause, however, seems to reflect just how far the authorities are prepared to accept the concept of university autonomy.

What is clearly less acceptable, however, is the prospect of further unrest. Before finally approving the NZS statutes, therefore, Dr Gorski made sure that they contained some limitation on future militancy. The ministry's revisions laid down that strike procedures would follow the pattern of the "Solidarity" statutes, with the additional proviso that a sit-in can be proclaimed only by the majority decision of a specially convened meeting of a students' organization, in consultation with the authorities of the higher educational establishment(s) concerned.

Vera Rich

US science budget

Reagan's way

Washington

In the absence of any explicit science policy — and, indeed, of any official presidential science adviser — the new Reagan Administration has adopted a "back-to-basics" approach in deciding where the budget axe should fall on federally sponsored research programmes.

Last week, Mr Reagan submitted to Congress the cuts which he is proposing, as part of a "national program for economic recovery", in the budget recommendations made by Mr Carter last month.

As expected, the main cuts will fall on those additional areas of responsibility which the federal government has taken on in recent years but which Mr Reagan is suggesting should be returned to individuals, to states or to private corporations and other institutions.

In the universities, for example, he proposes to eliminate a new fund for modernizing research equipment which Mr Carter had proposed creating in the National Science Foundation, with an initial budget of $75 million. Mr Reagan also wants to cut federal support for biomedical research institutions previously provided through the National Research Service Awards.

The Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) would see its budget cut by more than two-thirds,

0028-0836/81/090737-02$01.00

Threat to solar mission

Washington

European scientists could suffer directly from Mr Reagan's proposed budget cuts through a recommendation that the United States reduce its commitment to the international solar polar mission, a two-spacecraft programme run jointly by the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA).

Although no details of the reduced commitment have been officially announced, the White House is said to have suggested that NASA should stop work on its own spacecraft. This would effectively eliminate half of the project, and seriously reduce its scientific value, much of which depends on the simultaneous collection of data from the two spacecraft as they follow polar orbits around the Sun.

NASA and ESA officials are now discussing the implications of Mr Reagan's proposal, which could still be modified before the details of the new budget are announced on 10 March. The solar polar mission was threatened with termination by Congress last summer, but survived after vigorous intervention by the State Department and the Office of Science and Technology Policy.

Under the budget proposal submitted to Congress by President Carter last month, the two spacecraft would be launched from the space shuttle, using a modified Centaur launcher as a substitute for the delayed inertial upper stage, early in 1986. **David Dickson**

---

including an end to grants to states to protect their coastal zones, a 50 per cent reduction in support of college research under the Sea Grants programme, and the deferment of the National Ocean Satellite System (NOSS).

As for the private sector, a firm belief that the federal government should not interfere with the mechanics of the marketplace is reflected in substantial cuts to the Department of Energy's research and development budget. These would eliminate many of the department's efforts to demonstrate the commercial potential of new energy technologies, such as synthetic fuels, coal liquefaction and solar energy; but the aim is to maintain a basic commitment to long-range research projects considered too expensive or too risky by the private sector.

Public reaction to the proposed cuts from the scientific community has so far been muted. This is partly because precise details of where the cuts will fall will not be announced until 10 March and partly because there is little at present to be gained in Washington by speaking out against massive cuts in federal expenditure. Privately, however, laboratory chiefs and university presidents are already pulling all the strings they can, both within Washington's scientific establishment and among their congressional allies, to protect their own research programmes.

Some may already have been effective. A proposal from Mr David Stockman, director of the Office of Management and Budget, to eliminate the National Aeronautics and Space Administration's Galileo mission to Jupiter was removed from the President's message to Congress. Given a decision to defer the Venus Orbiting Imaging Radar, this would have virtually wiped out all future planetary research at NASA's Jet Propulsion Laboratory in California, Mr Reagan's home state.

On other proposals, there are bitter fights in prospect. Most of the projects that Mr Reagan is proposing to defer have been argued for by the scientific community strong and hard. These include NASA's gamma-ray observatory (already approved for funding by Congress), the National Science Foundation's 25-metre millimetre-wave radioastronomy dish and the NOSS satellite system.

Funding for space transportation systems will be maintained at a level adequate to cover the costs of the space shuttle, at the expense of slower development for Spacelab, and the rescheduling of space science flights — Galileo, for example, is likely to be shifted back from a 1985 to a 1986 launch. There will also be no funds for the solar electric propulsion system for which, in the absence of a Halley's comet mission, no applications have been approved.

At the National Science Foundation, budget restrictions will, as previously rumoured, be concentrated on programmes that are "narrowly focused or of less immediate priority" — such as innovation in small businesses and international scientific efforts — as well as on new initiatives in science and engineering education, and on research in the behavioural, social and economic sciences.

In contrast, there would be no reduction in the previously proposed 17 per cent increases for research in the mathematical and physical sciences, or the 20 per cent increase for engineering research. Both are considered by the new Administration to be "of relatively high importance to future technological advancement and to the long-term health of the nation".

In energy research, Mr Reagan is proposing a reduction of $40 million in the $607 million which had been suggested for basic energy sciences. Details of how this cut will be distributed are still being discussed. Most of it is likely to fall on high-energy physics, which accounts for two-thirds of the total, and will suffer the delayed construction of new facilities.

Biomedical research has been left relatively untouched ; where the previous Administration had suggested a relatively modest 9 per cent increase for the National Institutes of Health (NIH), Mr Reagan is suggesting a slight reduction in both 1981 and 1982 funding that will reduce this to 6 per cent. Much of the saving would come from reduced payment to educational institutions for NIH research training, which the Administration says would eliminate the practice of paying more to an institution for a federally supported trainee than would be charged for those who are not federally supported.

The Administration says that even though its proposed new budget for NIH would not fully cover the projected inflation rate — and that real reductions below the present base will therefore have to be made across all NIH institutes — it is committed to maintaining a substantial number of new research awards. It is therefore likely to continue the previous Administration's strategy of focusing on competitive project grants, rather than programme grants or intramural research.

David Dickson

UK nuclear energy

CEGB sheepish

The British nuclear industry is outwardly unruffled by criticisms from the House of Commons Select Committee on Energy last week (*Nature* 19 February, p.621). The general response is that the committee has misunderstood many of the issues. The strongest reaction is to the committee's criticism of the government's 1979 statement on nuclear power, which the committee took to be a commitment to build one nuclear station a year for the coming decade. This, it is said, was never the intention, so that the committee's recommendation that each reactor should be judged on its merits is already part of public policy.

The Central Electricity Generating Board (CEGB), the organization most sharply criticized, was the most sheepish last week. There is plainly some foundation for the charge that it had not been forthcoming with up-to-date estimates of cost and electricity demand. On the complaint that nuclear power stations are 34 per cent more expensive to build in Britain than elsewhere, Mr Glynn England, chairman of the board, is to meet nuclear suppliers and subcontractors to find ways of cutting costs and improving productivity on nuclear plant sites.

Mr England does not, however, accept the committee's view that the ordering of a second pressurized water station should be delayed for six or more years until the first, now being designed for the Sizewell site, is operating. Detailed studies and a public inquiry should provide enough information for the board to decide whether subsequent reactors should be based on pressurized water (as at Sizewell) or gas-cooled technology.

The Nuclear Installations Inspectorate also says that the committee has misunderstood its role. In particular, it

0028-0836/81/090738-02$01.00

says there would not be enough work for a full-time ultrasonics expert, recommended by the committee for testing pressure vessels, and that this and other work will be contracted out when it lacks the appropriate experts on its own staff.

The inspectorate is nevertheless still short of nuclear inspectors. Thirteen posts out of a total of 102 remain to be filled and salaries are a problem. Although nuclear inspectors' pay is as good as or better than that of other inspectors, it is below that in the industries from which it has to recruit. The problem — that the inspectors' salaries are linked to civil service pay — cannot be solved by removing the inspectorate from the Health and Safety Executive, which the inspectorate says would complicate licensing procedures.

The committee is said also to have misunderstood the role of the chief scientist at the Department of Energy and of the UK Atomic Energy Authority in advising the government on nuclear matters. The Department of Energy says that its chief scientist is responsible for nuclear advice but that there were exceptions when Dr Walter Marshall held the post as well as that of deputy chairman of the UK Atomic Energy Authority. Dr Marshall's version of this difficulty is different. He said last week that during his spell as chief scientist at the department, he felt no conflict of interest but, with the minister's agreement, meticulously kept the chairman of the authority informed of the advice he gave on nuclear matters.

The response so far to the report has been laconic and avoids detail. A more considered reply is likely to be published by the Department of Energy some months from now. Another contribution is likely to come from the Monopolies and Mergers Commission when it makes its views known on the structure of the electricity supply industry on 2 March.

Judy Redfearn

UK science research

On the move

The Science Research Council's attempt to foster mobility among British academics has made a modest beginning. The first four awards under the council's Special Replacement Scheme were announced last week. The scheme is designed to release senior academics from routine duties for five years, replacing them with younger people, usually postdoctoral researchers. The four awards will be followed by a further eleven before July, after which the council hopes to make ten awards a year.

The first four awards go to people who will be released from some or all of their administrative and teaching commitments for up to five years, enabling them to concentrate on their research interests. Each of the four departments is advertising a vacancy for a lecturer, whose appointment will be financed wholly by the Science Research Council for the first five years of his tenure. In every other respect, however, those appointed will be fully-fledged members of the academic staff.

Here the similarity ends. Some of the senior academics will return to their original position after five years. During that time, the department, which has guaranteed tenure to the new appointee, will have either found money elsewhere or lost a member of staff (by foul means or, more probably, fair). In other cases, the senior person will himself be retiring.

The Science Research Council, which has designed the scheme for flexibility, makes no stipulation about the areas of research involved, although it does intend to be represented on the selection boards for new appointees. In the cases so far announced, one professor should gain a member of staff in his own field of research, while another's department is to advertise for applicants for any of its research areas. Much discussion takes place behind the scenes between the council and the university concerned, and Sir Harry Pitt (ex-vice-chancellor of the University of Reading) is go-between and honest broker for the scheme.

The Science Research Council, through its boards and subcommittees which allocate grants among the applicants (30 of whom are now being considered), hopes to ease stagnation in research areas that it feels deserve encouragement. Awards have been given to Professors D. H. Everett (physical chemistry, Bristol), J. G. Powles (physics, Kent), M. Symons (chemistry, Leicester) and R. Butterfield (civil engineering, Southampton).

Philip Campbell

Authority more critical

The United Kingdom Atomic Energy Authority made a forceful response last week to the critical report of the House of Commons Select Committee on Energy (see *Nature* 19 February, p.621) Dr Walter Marshall, the bulky and voluble Welshman who succeeded Sir John Hill as chairman at the weekend, was quickly in action with his account of where the select committee had gone wrong.

The committee's wrath was directed chiefly at the Central Electricity Generating Board, but it also asked that the authority's role in the development of nuclear power in Britain should be restricted to research on long-term projects (fast reactors and fusion devices) and others where interested parties chose to commission work.

No-doubt Marshall

Marshall argues that this conclusion is mistaken. Thus he justifies the authority's work on the safety of pressurized water reactors (now costing £10 million a year) on the grounds that the authority is more independent than the would-be builders of the plant, the generating board, and less likely to continue indefinitely in the field than, say, the Nuclear Installations Inspectorate.

On the select committee's opinion that the authority should not be the public shareholder in the National Nuclear Corporation, the publicly supported construction consortium, Marshall says that the only effect of such a change would be to replace him by a civil servant as a director of the corporation. At present, he says, the authority's representation is the only source of independent criticism on the board.

The recommendation that the authority should quickly make an assessment of the Canadian CANDU heavy-water reactor system is similarly unwelcome. Marshall says that the committee has underestimated the difficulty of adapting even well-established reactor systems to British safety regulations, and estimates that a proper assessment would require two years of hard work. He points out that the select committee overlooked the authority's role in the development and management of the nuclear fuel cycle.

Dr Marshall's succession as chairman of the authority — he has been waiting in the wings for several years — presages a change of style. He is both outspoken and ebullient. He has his roots in the research establishment at Harwell (where he will keep an office). Last week he was saying that there will be no cause for changing the role of the authority until fast reactors are commercial realities some time in the next century. But he plans that the authority should become more skilled at explaining what it is about.

Community research

Project sharing?

Brussels

The European Parliament has now called for more community research. This arose at a meeting between the Parliament's Science and Energy Committee, the Dutch Minister for Science and Technology, Anton van Trier, and Dr Guenter Schuster, the director general for research, science and education of the European commission.

0028-0836/81/090739-02$01.00

Mr van Trier pointed out that between 1970 and 1980, Community expenditure on research and development had increased by 18 per cent, but that out of the 20,000 million European Units of Account (EUA) (1 EUA = £0.56) which the member states spent on their energy and research programmes in 1980, only 1.5 per cent was spent jointly. The president of the committee, Hanna Walz, while endorsing Mr Van Trier's affirmation of the significance of Community research, pointed out that there is a considerable gap between the declared intentions of the Council of Ministers and its actions.

Other members of the committee thought that the time was ripe to set in motion a "real Community research policy", since major areas could not be adequately tackled by each country on its own. Dr Schuster said that Community research would continue to concentrate on specific areas, particularly nuclear power, energy and environment, and pointed out that the Community contribution in these areas is 30 per cent, 10 per cent and 20 per cent respectively of total Community spending.

Mr van Trier and the committee agreed that a European market should be created for micro-electronic components and products, but that coordinating national programmes would only be feasible for large projects. The companies themselves must establish relations in order to coordinate the proposed research programmes. **Jasper Becker**

US Cabinet appointments

Conservatives in

Washington

One of the few bright spots for US environmentalists in the Cabinet appointments has been Mr Reagan's nomination of Mr James Buckley, a member of the US Senate from 1971 to 1977, as Under-Secretary of State for National Security, Science and Technology. If confirmed by the Senate, Mr Buckley will be directly responsible for the State Department's Bureau of Oceans and International Environmental and Scientific Affairs. Under a reorganization plan being discussed in the White House, the bureau may also be given responsibility for nuclear non-proliferation policy.

Mr Buckley is a conservative Republican. He voted against the War Powers Act giving Congress greater control over presidential decisions, and also opposed the first Strategic Arms Limitation Treaty (SALT), because it limited the United States to the construction of only two anti-ballistic missile sites — which were never built.

As a member of the Senate's environmental pollution subcommittee and its environmental science and technology subcommittee, Mr Buckley helped to forge much of the stringent environmental legislation of the 1970s. He clearly intends to take a close personal interest in environmental issues in his new post. High on his agenda is discussion with Canada of the effects of acid rain from the north-east states. Another issue is international attempts to prevent the extinction of endangered species; Mr Buckley told his nomination hearing in the Senate last week that "as a friend of the snail darter [whose potential extinction was used to hold up a federal dam project] I look forward to extending my responsibilities".

One potential dispute has already been headed off by the State Department. In its preliminary proposals for budget cuts, the Office of Management and Budget had suggested eliminating the $10 million contributed by the United States to the United Nations Environment Programme (UNEP). This — almost a third of the agency's budget — would have devastated UNEP's future programmes, but was successfully resisted by Mr Haig, and a contribution of $7.2 million is being proposed. In his Senate hearing, Mr Buckley said this would be "foolhardy".

Nuclear non-proliferation policy will be more difficult to deal with but will be high on the new Administration's agenda, since India is already preparing to make a further application for the export of enriched uranium to its Tarapur nuclear plant. Mr Buckley told the Senate Foreign Relations Committee that the Reagan Administration was completely reviewing non-proliferation strategy.

Mr Buckley, who ran for the Senate in Connecticut, had been tipped as the new administrator of the Environmental Protection Agency. In fact, the White House announced last week the nomination of a Colorado attorney, Mrs Anne Gorsuch, a close political ally of Interior Secretary James Watt, to head the agency. Both nominations have generated protests from environmentalists.

David Dickson

UK medical research

Money to spend

A £350,000 windfall has landed in the hands of the British Medical Research Council's Laboratory of Molecular Laboratory, Cambridge. The money was donated quite unexpectedly by Mr Thomas Usher, chairman of the Canadian company, Dextran. Since Mr Usher first approached the laboratory last May, a fund has been set up to support research fellows and students over the next seven years.

The offer of money comes with no strings attached, and the company had previously had only brief contact with the laboratory — many years ago Dr Max Perutz gave Mr Usher some advice.

The staff of the laboratory has agreed with Mr Usher that the money should be used to support research staff, the most serious hole left by cuts in the Medical Research Council's budget. The resulting fund will provide £50,000 a year for one to three-year fellowships, short fellowships for visiting scientists and studentships over the next seven years.

The scheme is to get under way in earnest next October. The trustees of the fund — Drs Sydney Brenner, director of the laboratory, Frederick Sanger, John Gurdon, Hugh Huxley and Aaron Klug — will decide on the next recipients in June.

Judy Redfearn

High-energy physics

China backs off

Washington

Economic problems have forced the People's Republic of China to postpone construction of a 50 GeV particle accelerator scheduled to be built near the Ming Tombs, north-west of Peking.

The accelerator, to have been built with the assistance of the US Department of Energy, was central to the agreement between China and the United States signed by President Carter and Vice-Premier Deng Xiao Ping in Washington two years ago.

US official who visited China last month were told that although construction was being deferred as part of a major cutback in capital construction projects, the delay did not imply a reduced commitment to research and training in advanced science and technology. The officials feel, however, that a "softening" in China's previous commitment to high-energy physics is inevitable.

Under the agreement, the Department of Energy would have carried out much of the construction of the components of the accelerator, receiving purchase orders over a period of five years of between $100 and $200 million for work that would have been carried out at the national energy laboratories. Subsequently the terms were changed so that the Chinese would carry out most of the design and construction work, and the United States would only provide basic training and advice.

The Chinese have also told the National Aeronautics and Space Administration that they will have to postpone the planned purchase of a new US-built telecommunications satellite system, another major component of the science and technology agreement. Four major contractors had already discussed with the Chinese the specifications of such a system, and were preparing bids. In a letter, the head of the Chinese academy of space technology, Dr Ran Xin-Mon, has now said that China was being forced to reconsider the whole of its space programme in the light of its economic situation.

China is, however, continuing negotiations over the purchase of equipment to receive geophysical information from the LANDSAT remote sensing satellite. It has

0028-0836/81/090740-02$01.00

also said that it intends to reschedule the satellite deal when funds are available.

According to Representative Donald Fugua, chairman of the House of Representatives Science and Technology Committee, who headed the delegation to Peking, China has decided that it will have to cut out about half of its ambitious construction plans in science and technology — including plans for a new steel mill to have been built outside Shanghai. But Chinese officials have stressed that other activities, such as scientific exchanges and research programmes will not be affected.

David Dickson

Soviet space research

Signs of strain

The planned Franco–Soviet mission to Halley's comet in 1986 turns out to involve a substantial cutback for the planned joint mission to Venus in 1984. This development, first indicated last summer, was finally confirmed two weeks ago in Paris. Two of the original four Venus probes have now been assigned to Halley, so that the French have had to abandon plans for a large balloon which would have analysed the atmosphere of Venus. Although the reduced Venus mission could still have accommodated a smaller balloon, the consequent problems of high-temperature electronics could not be solved in time.

According to M. Hubert Curien, president of the Centre National d'Etudes Spatiales, the French side has few regrets. The new programme, he said, is "very full and very attractive". Nevertheless, the final announcement of the scaled-down Venus mission was a reminder that the resources of the Soviet space programme, though vast, are not infinite. It came only a few days after *Pravda* had published two long articles of no great topicality clearly intended to present the Soviet space effort in a favourable light.

One dealt with the detection of iron in the lunar regolith samples recorded by the Luna-16 probe in 1970 and one with the seventieth anniversary of the birth of the late Mstyslav Keldysh, whom Khruschev brought out of the field of military rocketry to lead the academic space programme.

The question of the scale of Soviet space research will arise this week, when the twenty-sixth congress of the Communist party of the Soviet Union will be required to approve the basic guidelines for the next Five-Year Plan. As with the two previous plans, these will include a commitment to space exploration "in the interests of the national economy". This is an empty formula for deep-space research, although satellite photography is playing an increasing part in a number of aspects of Soviet planning, from fish-spotting to geological surveying. The costs of the programme are, however, never published.

Hints that Soviet space spending may be subject to increasing financial constraints have, however, been dropped in recent months. There seem to be no further plans for Comecon participation in manned flights after Mongolia and Romania have put a cosmonaut in orbit. Soviet planners have so far failed to respond to Bulgarian hints that, as their cosmonaut, Georgi Ivanov, failed to complete his mission (his Soyuz transport craft could not link up with the Salyut station), Bulgaria should be allowed another turn, especially in its 1,300th year of statehood. Instead, Bulgaria has been promised two unmanned probes instead of the original one.

The Soviet commitment to Comecon participation in space research nevertheless continues. A new scientific cooperation programme with Poland, announced last month, put special emphasis on space research. Poland, the homeland of Copernicus, may have a special place in Soviet space planning but the Soviet Union's contribution to Comecon collaboration is substantial. It pays the total cost of the launching and ground control. The participating Comecon partner simply has to pay for its own apparatus and the associated data processing — a privilege also extended to the French.

Vera Rich

Science and government

Lords look now

The House of Lords Select Committee on Science and Technology has taken the unusual step of making a public appeal for opinions on the subject of its latest inquiry — science and government. The inquiry's chairman will be Lord Sherfield. Lord Todd, who is thought to have instigated the inquiry, and who was expected to take the chair, seems to have stepped down in the belief that his strong views can be better aired from the body of the committee.

Central to the inquiry will be the need for a chief scientific adviser to the government and the success of the Rothschild customer-contractor principle. The chief scientific adviser's post was abandoned in 1974 after the Rothschild report recommended that individual government departments should take more responsibility for seeking advice and commissioning research. The system of departmental chief scientists that resulted was intended to enable government departments (the "customers") to commission research within the research councils (the "contractors") with money transferred to them from research council budgets.

Although the principle has worked well in some departments, it has been disastrous in others. The Department of Health and Social Security has acknowledged that it cannot place contracts for biomedical research, while the Ministry of Agriculture, Fisheries and Food is considering a proposal to abolish the post of departmental chief scientist as part of its changing relationship with the Agricultural Research Council (*Nature* 1/8 January, p.2).

The committee started taking oral evidence yesterday (25 February) from Sir Ian Bancroft, head of the home civil service. Next on the list are Lord Trend, former Secretary to the Cabinet and author of the 1964 report on the organization of civil science, and Sir Hermann Bondi, chairman of the Natural Environment Research Council and a former chief scientist at the Ministry of Defence and Department of Energy. Those wishing to submit evidence should write to the Clerk of the Select Committee on Science and Technology, Committee Office, House of Lords, London SW1 by 31 March 1981.

Judy Redfearn

Questions to answer

A *Machinery of government*

(1) Should scientific advice and/or research procurement be a distinct function of government separate from the existing departmental structure?

(2) How successful is the system of departmental chief scientists in procuring advice, managing research and influencing policy?

(3) How well supported are ministers when judging scientific priorities in decision making, particularly if government departments are not in agreement?

(4) How far is the scientific advice sought by government geared to supporting predetermined objectives?

B *Finance*

(1) How satisfactory is the division of financial responsibility between the research councils (as a group) and government departments funding research on the customer-contractor principle?

(2) Is any research which could be of real value to government being neglected for lack of identified customers or because it is peripheral to the interest of several customers; if so, what changes could rectify this?

(3) Are any changes in research budgets desirable?

C *Machinery of science*

(1) How adequate are the channels of communication from the scientific community to government, and vice versa?

(2) Is there satisfactory contact between those administering science in higher education, industry, the research councils and government?

(3) How could statutory procedures for consultation by government in scientific matters be improved?

(4) Are existing sources of advice adequate to ensure that the United Kingdom gains all it can from EEC and international research programmes?

D *Scientific manpower*

What manpower constraints are there on the provision of scientific advice to government?

0028-0836/81/090741-01$01.00

CORRESPONDENCE

Genes and race

Sir — I write in reply to Dr Steven Rose's letter (*Nature* 22 January, p.335), which drew attention to the fact that a right-wing journal, *New Nation*, has quoted me, together with other evolutionary biologists, in support of their view that our genetic constitution makes it impossible for us to live in a racially integrated society. I welcome the opportunity to say that there is nothing in modern evolutionary biology which leads to this conclusion.

John Maynard Smith

University of Sussex,
Falmer, UK

Museum debate

Sir — I have been following the "great museum debate" in your pages with a profound sense of detached amusement. But as matters are quickly reaching a level of absurdity that may inspire me to write the 15th Gilbert and Sullivan opera, and as I am, in a sense, the focal point for Halstead's glorious uproarious misunderstanding, I suppose I should have my say.

Halstead began all this by charging that the venerable Natural History Museum is now purveying Marxist ideology by presenting cladism in its exhibition halls. The charge is based on two contentions: (1) a supposed link between the theory of punctuated equilibrium, proposed by Niles Eldredge and myself, and cladistic philosophies of classification; and (2) an argument, simply silly beyond words, that punctuated equilibrium, because it advocates rapid changes in evolution, is a Marxist plot. For the first, there is no necessary link unless I am an inconsistent fool; for I, the co-author of punctuated equilibrium, am not a cladist (and Eldredge, by the way, is not a Marxist, whatever that label means, as if it mattered). Under cladism, branching events may proceed as slowly as the imperceptible phyletic transitions advocated by the old school. Punctuated equilibrium does accept branching as the primary mode of evolution, but it is, fundamentally, a theory about the characteristic rate of such branching — an issue which cladism does not address.

For Halstead's second charge, I did not develop the theory of punctuated equilibrium as part of a sinister plot to foment world revolution, but rather as an attempt to resolve the oldest empirical dilemma impeding an integration of palaeontology into modern evolutionary thought: the phenomena of stasis within successful fossil species, and abrupt replacement by descendants. I did briefly discuss the congeniality of punctuational change and Marxist thought (*Paleobiology*, 1977, p.145) but only to illustrate that all science, as historians know so well and scientists hate to admit, is socially embedded. I couldn't very well charge that gradualists reflected the politics of their time and then claim that I had discovered unsullied truth. But surely Halstead, who has done some statistics in his day, knows that correlation is not cause. If I may make a serious point: I grew up frightened in a leftist household during the worst days of McCarthyism in America; and I know that what seems peripheral or cranky today can become a weapon tomorrow (consider the current creationist surge in America). May we avoid red-baiting; it may not always be harmless.

I saw the cladistic exhibits last December. I did not care for them. I found them one-sided and simplistic, but surely not evil or nefarious. I also felt, as a Victorian aficionado who pays homage to St Pancras on every visit to London, that most of the newer exhibits are working against, rather than with, the magnificent interior that houses them. But I would not envelop these complaints in ideological hyperbole; Halstead has said enough.

Stephen Jay Gould

Museum of Comparative Zoology,
Harvard University,
Cambridge, Massachusetts, USA

Last word?

Sir — It is a little late in the day for Stephen Gould to try and come the innocent. If he did not want "Halstead's glorious uproarious misunderstanding" to get under way, he should have avoided dragging Engels on to his side in the first place[1]. Mind you, even this might not have worked, for the "political" implications of punctuated equilibria have not gone unremarked, as some of the recent correspondence has made manifest.

I have taken Engels and Lenin as my main sources on dialectical materialism, which I have sought to apply in the conduct of my own researches. There is only one point, but this is fundamental, at which I would part company with Marxism and that is the nature of qualitative changes, which I see not as sudden leaps but as gradual in the tradition of Charles Darwin. I have recently been nonplussed to learn from Stephen Gould that "many orthodox Marxists have been quite content with Darwinian gradualism" (ref. 2). This is equivalent to someone, who insists that Christ was a myth, being considered an orthodox Christian. But perhaps in the United States, I would be deemed an orthodox Marxist!

When it comes to the cladists with their punctuated tendencies, we run into a perfectly ludicrous source of semantic confusion. If, as some of the correspondents have insisted, cladistics is concerned only with pattern and not process, then obviously there is no point in arguing further, because I am concerned primarily with process. I wrote about the kind of classical Hennigian cladistics actually being presented in the public galleries of the British Museum (Natural History) and clearly explained in their accompanying booklets, and not the new transformed variety of Patterson[3] and others.

Tempting though it may be, I am sufficiently modest to decline the mantle of oracle proffered by Rosen[4] with regard to the origin of *Homo sapiens* from *Homo erectus*. I claim no special insight in these matters but merely reported the considered and published consensus of the staff of the British Museum's own Sub-Department of Anthropology. Wood[5] has drawn attention to the "dubious academic practice" of ignoring "uncomfortable" evidence — the scandal of this is that it was deliberate, involving, as it did, the overruling of the museum's own experts. Critical scientific evidence is being deliberately withheld from the public who are, in consequence, being seriously misled as to the nature of "Man's place in evolution". All is apparently being subjugated to a chosen dogma. I have been roundly abused for implying that a far left political connection might be involved, but the present dogmatic policies seem to bear its unmistakable stamp.

The British Museum (Natural History), London, is a major public scientific institution in this country and as such should be accountable to the public. It is surely reasonable to expect the Director to answer the charges that have been levelled — a continued "dignified silence" in the present circumstances is simply not good enough.

L. B. Halstead

Departments of Geology and Zoology,
University of Reading, UK

1. Gould, S. J. & Eldridge, N. *Paleobiology* **3**, 225-251 (1977).
2. Gould, S. J. (personal communication).
3. Patterson, C. *Nature* **288**, 430 (1980).
4. Rosen, D. E. *Nature* **289**, 8, 105 (1981).
5. Wood, B. A. *Nature* **289**, 8 (1981).

Badger controversy

Sir — On 11 December 1980 the Mammal Society published a letter in *Nature* drawing attention to some of the many anomalies in the Zuckerman report. The reasons we started an open discussion of the report are obvious. Badger gassing started in 1975; nearly six years later, after the expenditure of a great deal of time, money and effort, it is our belief that there is no unequivocal evidence that the present gassing policy is likely to produce a long-term solution to the problem. We suggested that the Zuckerman report is one-sided in its interpretation of the evidence, and that the conclusions and recommendations in the report are too categorical and do not take sufficient account of the complexity of the problem.

Following our original letter, three letters have been published in *Nature*. Dr Plowright (1/8 January 1981, p.8) presented no new information, nor did he answer any of the points we raised, and so his letter will not be discussed further. Dr Yates (22 January 1981, p.218) questioned a graph sent to Lord Zuckerman for his comments. This graph was simply intended to show that the rate of decline of TB was similar in the South West to that elsewhere in England. That the incidence of TB was different in the two areas was never disputed. Dr Yates published an alternative graph, which was simply another presentation of the same data; it showed that the incidence of TB varies in different parts of the South West. That is the very point we stressed in our original letter. Dr Yates' graph also showed that there was a decline in TB in all the regions sampled, irrespective of whether badger gassing was carried out in that area. That is the only point our original graph was designed to demonstrate.

In fact Dr Yates' graph has highlighted the

Continued on page 831

0028-0836/81/090742-02$01.00

NEWS AND VIEWS

Models of speciation — the evidence from *Drosophila*

Library
College
Upper Circular Road
CALCUTTA

from J.S. Jones

If Lewontin's (1974)[1] claim that "we know almost nothing of the genetic changes that occur in species formation" is still true then we must by now know it in remarkable detail, for the literature of evolutionary biology has been filled with speculations, observations and experiments on this very subject for the last five years. There are two models of the origin of species. In one of these[2] the division of a population into two genetically incompatible entities is seen as a by-product of the gradual adaptation of its geographically isolated parts to different selection pressures. Hybrids between the evolving populations are relatively unfit because they have accumulated different alleles with individually minor effects. Reproductive isolation thus arises at first from the incompatibility of separately developing gene pools; should these pools later come into contact, selection may lead to the evolution of behavioural or ecological barriers which reduce the frequency of hybrid matings. The origin of a new species is the result, writ large, of the familiar process of natural selection within an existing species.

A second model — that speciation is brought about by a 'genetic revolution' — has recently returned to prominence[3-5]. This theory places less weight on selection, and emphasizes the importance of random processes in the origin of taxonomic groups. A small and isolated population of an originally widespread and highly variable species will, it is suggested, undergo major genetic changes because of the chance loss of alleles in the small founding sample, the increased frequency of homozygotes caused by inbreeding in the early generations and the resulting disruption of interacting complexes of genes and of developmental processes under genetic control. The new isolate may move into previously unexploited ecological niches and undergo an enormous increase in numbers (the 'population flush'[3]) with a consequent relaxation of selection against deviant individuals which will further favour the success of the genetic revolution. Rapid fluctuations in population size and random gene frequency changes in these early generations will quickly lead to the evolution of reproductive incompatibility with the parental population. According to this view, isolation causes speciation, rather than sometimes allowing it to take place as an incidental result of slow adaptive evolution.

Recently, some palaeontologists have claimed that this model of speciation can explain the apparent gaps in the fossil record[6]. Rather than accepting that the gaps reflect the record's incompleteness, they claim that they are the result of a pattern of bursts of rapid speciation interspersed with long periods of evolutionary stability. Only rarely will members of the small speciating population be fossilized, so that the lineage will show what are known as 'punctuated equilibria' — a long series of unaltered fossil forms followed by a rapid shift to a new type. Species are therefore discrete units not only in space, but also in time.

Some biologists have interpreted this view of evolution as an attack on the fundamentals of Darwinism and a bitter dispute is still taking place. However, the problem can be approached experimentally using *Drosophila*. There is now good evidence that some populations may indeed contain hidden genetic variation sufficient to cause 'instant speciation' and that rapid fluctuations in population size can sometimes produce reorganizations sufficient to lead to reproductive isolation within a very few generations.

Drosophila mercatorum is normally sexual, but in the laboratory a female can produce parthenogenetic clones. It is possible to freeze the recombinant offspring of a single heterozygous female into a number of diverse clones, and by comparing these to estimate the amount of hidden genetic variation in the original population. Templeton and his colleagues[7,8] find that such clones differ from each other in morphology, in behaviour and in aspects of fitness sufficient to cause rapid changes in their frequency when competing in a population cage. Most important, when females from some of the derived clones are tested for mating potential with males from their ancestral sexual population, they show strong reproductive isolation from them. What is effectively a new species of *Drosophila* has therefore arisen as a result of reorganizations of the genome of a population within a few generations.

Sexual populations may also undergo genomic transformations of this kind. A population of *D. pseudoobscura* was placed in a large cage and allowed to expand rapidly in numbers[9]. At the peak of the population flush, 8 single pairs were chosen to start 8 new populations, each of which grew to a large size. These populations were put through three further cycles of reduction and expansion. Interpopulation tests of mating preference showed that in a third of the derived cages there was a significant tendency for individuals to prefer to mate with others from their own population. Reproductive isolation was the result of the demographic (and associated genetic) changes within the populations, as control populations isolated from each other without undergoing rapid changes in population size did not mate preferentially.

Population flushes in nature are most likely to occur in organisms which disperse easily, have short generation times and live in regions where new and unexploited habitats become available. The *Drosophila* of Hawaii are a good example of such a group. The Hawaiian archipelago came into being some 40 million years ago, and the youngest island is about 700,000 years old. Nearly a third of all known *Drosophila* species occur only in Hawaii, suggesting that speciation has been extremely rapid. Carson[10] suggests that this reflects a history of founder events, population flushes, and the formation of new gene combinations as flies blew from one island to that formed next in the volcanic chain. On the basis of chromosome phylogeny he estimates that as few as 22 inter-island colonizations (many of which may have involved just one fertilized female) over only 6 million years will account for speciation in the best known group of Hawaiian *Drosophila*. The rapidity of speciation is a result of random events and demographic changes leading to reproductive isolation, rather than of slow adaptive evolution on each island.

The patterns of reproductive isolation between flies from different islands support this view[11]. In *Drosophila* the female selects the mate; a courting male must produce the correct sequence of behaviours before being accepted, but a male will attempt to mate with almost any available female. Reproductive isolation between species of Hawaiian *Drosophila* from different islands is often asymmetrical. A female from a new species on a young island will accept a male from the ancestral population, but a male of the new species cannot successfully court a female of the ancestral species. This is evidence for

---

J.S. Jones is in the Department of Genetics and Biometry, University College London.

0028-0836/81/090743-02$01.00

rapid speciation in a founding population; the males of the derived species lost, by chance, in their small founding population some of the behavioural repertoire which made them acceptable to ancestral females, but the ancestral males retain all the genes for the behaviour needed to stimulate females of both species.

This process has been simulated in the laboratory by an inadvertent experiment of the kind familiar to *Drosophila* geneticists[12]. A stock of Hawaiian *D. silvestris* underwent at least two population crashes (perhaps to a single fertilized female) during its stay in the laboratory. The derived stock showed strong behavioural isolation from freshly collected *D. silvestris* from the same location. Wild males successfully courted derived females, but derived males could not stimulate females from the ancestral population. Asymmetrical reproductive isolation did not develop as a by-product of slow adaptive evolution but resulted from extreme demographic changes.

As species are, by definition, genetically isolated from each other, it might seem impossible to investigate the genetics of reproductive isolation by carrying out breeding experiments. This apparent contradiction has been overcome in the Hawaiian *Drosophila*[13]. *D. silvestris* and *D. heteroneura* are known to be closely related, but differ markedly from each other in head shape. This is an effective behavioural isolating mechanism in the wild. It is possible to make fully fertile crosses between these species in the laboratory, and genetic analysis shows that the differences in head shape are due to a single sex-linked locus which influences the phenotypic expression of about 10 interacting autosomal loci. The genetic basis of each species' integrity thus depends on rather few loci — a result consistent with theories emphasizing the importance of random events in the origin of species.

Although these views of speciation as a 'revolutionary' process are now receiving much attention, there is also evidence strongly supporting the view that many species originate through the slow evolutionary accumulation of small variations. Recent theoretical work by Lande[14] shows just how effective selection acting on many genes with individually small effects can be in leading to the divergence of large populations. His predictions are supported by many experiments in which individual *Drosophila* prefer to mate with others from the same natural population[15] or in which the offspring of interpopulation crosses are less fit than those whose parents arise from the same population[16]. Laboratory stocks subjected to different temperature and humidity stresses accumulate over a period of time genetic differences great enough to lead to mating preferences or hybrid inviability[17]. Reproductive isolation has also been found to occur between populations of *Drosophila* and of house flies selected for high or low activity, for different bristle number and for DDT resistance[18-21].

In summary, there is convincing evidence for both models of the origin of species. In some lineages, speciation may result mainly from natural selection acting on genes with small effects, while in others it probably results from the disruption of a genetic system following the random colonization of a new environment by a small population. The two mechanisms are not mutually exclusive and within one genus both may be important. In *D. mercatorum* (where experiments with derived clones show that natural populations contain the genetic potential for very rapid speciation) geographically separated populations have differences in their behaviour which have arisen by the accumulation of many genes with small effects — and which are sufficient to cause considerable reproductive isolation[22]. In the Hawaiian *Drosophila*, also, where evidence for rapid speciation as a result of random effects is so strong, some populations from the same island have undergone considerable divergence which is not a result of founder events[23]. Any complete theory of the origin of species must recognize — as did Darwin himself — the importance of the joint action of random processes and of natural selection.

Models of speciation have recently become entangled with claims that the 'punctuated equilibrium' theory of evolution favoured by those responsible for the dinosaur gallery in the Natural History Museum reflects that of the evolution of society held by Karl Marx[24]. Whatever the scientific merit of an argument couched in terms of Dinosaurs versus Democracy, it is worth considering the attributes of a lineage liable to undergo rapid speciation by genetic revolution[4,5]. They include an abundant and freely interbreeding ancestral population, high individual mobility, and rapid reproduction. Although the *Drosophila* of oceanic islands fit these criteria well, it is hard to imagine any group less able to satisfy them than the dinosaurs. In particular, the idea of a windblown propagule of a pregnant female *Brontosaurus* is enough to strain the imagination of even a palaeontologist! ☐

---

1. Lewontin, R.C. *The Genetic Basis of Evolutionary Change* (Columbia, 1974).
2. Mayr, E. *Animal Species and Evolution* (Harvard U.P., 1963).
3. Carson, H.L. in *Population Biology and Evolution* (ed. R.C. Lewontin, 1968).
4. Templeton, A.R. *Genetics* **94**, 1011 (1980).
5. Templeton, A.R. *Evolution* **34**, 719 (1980).
6. Eldridge, N. & Gould, S.J. *Models in Paleobiology* (ed. T.J.M. Schopf, Freeman, 1972).
7. Annest, J.L. & Templeton, A.R. *Genetics* **89**, 193 (1978).
8. Templeton, A.R. *Genetics* **92**, 1265 (1979).
9. Powell, J.R. *Evolution* **32**, 465 (1978).
10. Carson, H.L. & Kaneshiro, K. *A Rev. Ecol. Syst.* **7**, 311 (1976).
11. Kaneshiro, K. *Evolution* **34**, 437 (1980).
12. Ahearn, J.N. *Experientia* **36**, 63 (1980).
13. Templeton, A.R. *Evolution* **31**, 630 (1977).
14. Lande, R. *Am. Nat.* **116**, 463 (1980).
15. Ayala, F.J. *Evol. Biol.* **8**, 1 (1975).
16. Ehrman, L. & Parsons, P.A. *Behav. Genet.* **10**, 401 (1980).
17. Kilias, G., Alahiotis, S.N. & Pelecanos, M. *Evolution* **34**, 730 (1980).
18. Van Dijken, F.R. & Scharloo, W. *Behav. Genet.* **9**, 555 (1980).
19. Hurd, L.E. & Eisenberg, R.M. *Am. Nat.* **109**, 353 (1975).
20. King, J.C. *Cold Spring Harb. Symp. quant. Biol.* **20**, 311 (1955).
21. Thoday, J.M. *Proc. R. Soc.* B **182**, 109 (1972).
22. Ikeda, H. Takabatake, I. & Sawada, N. *Behav. Genet.* **10**, 361 (1980).
23. Craddock, E.M. & Johnson, W.E. *Evolution* **33**, 137 (1979).
24. Halstead, L.B. *Nature* **288**, 208 (1980).

---

100 years ago

MIGRATION OF THE WAGTAIL

I fear I may be attempting to trespass too frequently on the columns of *Nature* recently, but the paper on the subject of wagtails taking a passage on the backs of cranes in a long flight, resembles so much a somewhat similar story told and believed in by the Indians in several parts of North America, that I venture to send you an account of it.

All the Indians (Maskegon Crees) round the south-western part of Hudson's Bay, assert that a small bird of the Fringillidae tribe takes a passage northward in the spring on the back of the Canada goose (*A. Canadensis*), which reaches the shores of Hudson's Bay about the last week of April.

They say that they have often seen little birds fly away from geese when the latter have been shot or shot at.

An intelligent, truthful, and educated Indian named George Rivers, who was very frequently my shooting companion for some years, assured me that he had witnessed this, and I believe I once saw it occur.

It is only the Canada goose that these little migrants use as an aerial conveyance, and certainly they both arrive at the same date, which is a week or two earlier than the other kinds of geese (*A. hyperboreus* and *albifrons*) make their appearance. I knew the little bird well, but it is so long ago that I have forgotten the name.

The Indians on the shores of Athabasca and Great Slave Lakes, both great resorts of wild geese, tell a similar story. If a fabrication, I do not see why it should be invented about the Canada goose only, and not about other species which are equally numerous.

All the Coast Indians of Hudson's Bay devote a month or more every spring to wild fowl (chiefly geese) shooting. As soon as the geese begin to arrive, the Indian constructs a concealment of willows and grass, usually near a pool of open water, at the edge of which he sets up decoys. When geese are seen approaching (usually flying at a great height) the Indian imitates their call, and the geese on seeing the decoys circle round, gradually coming lower down until within shot, when they are fired at. It is from these high-flying geese that the small birds are seen to come.

If the geese are flying low it is a pretty sure indication that they have already rested on the ground somewhere near, after their long flight, when of course their tiny passengers would have alighted. JOHN RAE

Royal Institution, February 26

from *Nature* **23**, 3 March, 411, 1880

0028-0836/81/090744-01$01.00

Last supper of two Indian Triassic reptiles?

from Barry Cox

Life restorations of two specimens of *Malerisaurus*, showing quadrupedal and bipedal gaits.

DURING the winter of 1965–66, Indian palaeontologists found a pair of skeletons of the crocodile-like phytosaurian reptiles in the late Triassic rocks of the Maleri Formation in the Pranhita-Godavari Valley, Andhra Pradesh, central India. They were found so closely associated that they soon became known as 'the loving couple'. During the removal of the soft, clay-like matrix from their skeletons, it soon became clear that they had also eaten a similar last meal, for within the rib cage of each there appeared the skeleton of a smaller reptile.

These two skeletons have now been described by Chatterjee (*Phil. Trans. R. Soc.* **B291**, 163; 1980) as a new genus, *Malerisaurus*. It belongs to a group of still poorly known and poorly understood Permo-Triassic lizard-like reptiles that includes the ancestors of the lizards themselves. *Malerisaurus* was of considerable size, some 1.3m long, and it is notable for its rather large hind-limbs. The length of the fore-limb is only 59 per cent of the length of the hind-limb, and the latter is 74 per cent of the length of the trunk. Proportions of this kind are found today in a number of facultatively bipedal lizards, such as the frilled lizard *Chlamydosaurus*, and the life restorations of the two specimens that illustrate Chatterjee's paper show them in two alternative poses. The dentition of *Malerisaurus* suggests that it fed on the general range of invertebrates that is described as an 'insectivorous diet'.

The sedimentology of the Maleri deposits suggests that the environment of this part of India during the late Triassic was probably very similar to that found there today, with a subtropical monsoon climate. *Malerisaurus* therefore probably lived in a densely forested, lowland environment with lakes or flood-plains.

The Maleri fauna also contains a variety of other reptiles, as well as the large aquatic amphibian *Metoposaurus*. All of these belong to groups that are widely known throughout the Triassic world, and attest to the ease of dispersal through that ancient Pangaea world-continent.

Barry Cox is Professor of Zoology at Kings' College, London.

Is the Universe expanding?

from David A. Hanes

THE DISCOVERY by Edwin Hubble[1] that the Universe is not static, but an evolving entity, is surely the most important finding in extragalactic astronomy this century. Few astronomers would contest the view that the Universe is in a state of general smooth expansion. However, recent provocative statistical studies by J. F. Nicoll, D. Johnson, I. E. Segal and W. Segal[2] of the relationship between the redshifts of moderately nearby galaxies and their apparent brightnesses have brought this simple yet profound picture under attack.

In a general expansion, the recession velocity of a galaxy (or empirically its Doppler redshift) should be directly proportional to the distance of that galaxy. Nicoll *et al.* claim they can reject such a relationship; rather, their findings imply a quadratic relationship between redshift and distance for a moderately nearby sample of galaxies (those with recession velocities less than 5,000 km s^{-1}). Such a discovery is consistent with the predictions of I.E. Segal's[3] alternative world model, the chronometric cosmology, and at variance with the predictions of Einsteinian General Relativity. At the very least, these statistical demonstrations demand explanations in terms of unsuspected selection effects, subtly biased samples, or local anisotropies or inhomogeneities; and, should convincing explanations be lacking, they may suggest that the usual cosmological descriptions of the Universe are inadequate.

Edwin Hubble's original data suggested a linear relationship between velocity (inferred from redshifts) and distance (inferred from apparent luminosities) for galaxies. This relationship has an appealing simplicity, for in a universe in which we are in no preferred position and where there are no preferred directions — that is, in a universe where the 'cosmological principle' holds — it can be shown that such uniform expansion (or contraction) is the only permitted large-scale motion. The cosmological principle is justified both philosophically and empirically: the former in the neo-Copernican sense that we must assume that we are in no way singled out as special observers of the cosmos; and the latter in that galaxy counts and the isotropy of the cosmic background radiation suggest that the Universe is indeed homogeneous and isotropic on sufficiently large scales[4]. There are other reasons for the faith of astronomers in the expansion hypothesis: the expansion is associated with a characteristic time (which may naively be thought of as the time since the Universe was in a state of infinite density) which is closely comparable with the ages of the oldest known stars; and the discovery[5] of the predicted[6] cosmic microwave background strongly suggests that the Universe has evolved from a hotter dense state. Moreover, the Hubble expansion finds a natural theoretical framework — that of Einsteinian General Relativity — which in fact predates[7] the discovery of the expansion and which has been amply confirmed in the limited tests possible to date[8].

For all these reasons astronomers will be unmoved, except by unequivocal tests, to adopt alternative cosmological descriptions of the Universe. Segal's chronometric cosmology is one such description. It has been characterized[9] as a 'gravitational tired light' theory; that is, one in which observed redshifts for remote galaxies arise through the loss of energy by photons during their propagation rather than because of any systematic motions. In Segal's

David A. Hanes is at the Anglo-Australian Observatory in New South Wales, Australia.

0028-0836/81/090745-02$01.00

formulation, the chronometric theory reproduces many of the observational attributes of the cosmos, such as the universal background radiation; and the telling coincidence of inferred expansion age with stellar age is not a problem[10], although arbitrarily old galaxies may abound. Astronomers — indeed, all scientists — are reluctant to adopt purely *ad hoc* theories, those which follow rather than predict fundamental discoveries and which explain later observation by judicious manipulation of the theory, and this kind of criticism has been levelled at the now largely discredited Steady State cosmologies. However, Segal's chronometric cosmology does make a crucial prediction relating to the behaviour of the redshift–apparent magnitude relation for systems of moderate redshift (inferred velocities $\lesssim 5{,}000$ km s^{-1}): the observed relation should be quadratic.

The statistical study of such nearby systems is plagued by a variety of important selection effects. Complete redshift determinations are available only for the brightest few hundred galaxies. At fainter levels, galaxies of special interest — those in rich clusters, or those associated with bright galaxies — may provide additional data. The galaxy luminosity function is broad, and an apparently bright galaxy with well determined redshift may be a remote very luminous object, while nearby fainter galaxies go unstudied; thus volume-limited samples are not to be hoped for. Perhaps most important, the local dynamics enter into any analysis: the corrections for solar motion around our own galactic centre, the peculiar motions of single galaxies and the dispersion of velocities because of interactions in clusters of galaxies are all relatively large compared with Hubble expansion rate for the nearby systems: the motion of our own local group of galaxies, which appears to be falling into the rich Virgo cluster[11], may introduce systematic effects; and other anisotropies may be present.

Deviations from the simple Hubble law have in fact been noted in the past. Hubble[13] himself recognized a degree of nonlinearity in his data and concluded that the gravitational deceleration of the expansion was the cause. Vaucouleurs[13] interpreted similar findings in terms of local departures from the Hubble flow caused by the existence of a Local Supercluster of galaxies. Other anisotropic features in the flow, the so-called Rubin–Ford effect, have received various suggested explanations[14]. The lesson to be drawn is that the purely astronomical interpretation of such effects in not at all clearcut, even in the simplifying context of a general underlying expansion.

Statistical tests must obviate these problems by the careful definition of completely unbiased samples unaffected by subtle selection effects. The difficulty of the exercise is reflected in the fact that R. M. Soneira[15], in a careful statistical reworking of Segal's earlier analysis, found no disagreement with the expansion hypothesis but felt able to exclude the correctness of the chronometric prediction at better than the 99.999 per cent formal confidence level. However, Soneira's analysis has in turn been criticized by Segal[16] and by Nicoll and Segal[17], who reach the opposite conclusion at comparable confidence levels. Their latest papers represent the most exhaustive statistical analysis to date, with the authors concluding that. . . "the Hubble law lacks an objective statistical foundation".

The validity of this conclusion and the resolution of the questions raised about our understanding of the large-scale structure of the cosmos may come in further statistical analyses which confirm or deny the various treatments. It may be that considerably more observational data and careful mapping of local anisotropic effects will prove essential. If the suggested nonlinearities are found to exist, most astronomers will seek explanations in terms of local deviations from the Hubble flow rather than in the wholesale rejection of the presently believed cosmological description of the Universe, with its history of successful predictions and supportive tests. The onus is upon the advocates of alternative theories to demonstrate the absolute insufficiency of any such explanations. □

1. Hubble, E. *Proc. natn. Acad. Sci. U.S.A.* **15**, 168 (1929).
2. Nicoll, J.F., Johnson, D., Segal, I.E. & Segal, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6275 (1980).
3. Segal, I.E. *Mathematical Cosmology and Extragalactic Astronomy* (Academic, New York, 1976).
4. Peebles, P.J.E. *Physical Cosmology*, 26 (Princeton University Press, 1971).
5. Penzias, A.A. & Wilson, R.W. *Astrophys. J.* **142**, 419 (1965).
6. Gamow, G. *Phys. Rev.* **74**, 505 (1948).
7. Einstein, A. *Sitzungs-Berichte der Prussischen Akad. der Wissenschaften* (1917).
8. Shakeshaft, J.R. *Observatory* **99**, 122 (1979).
9. Pecker, J.-C. *IAU Colloq.* No. 37, 451 (1977).
10. Segal, I.E. *IAU Colloq.* No. 37, 603 (1977).
11. Cheng, E.S., Saulson, P.R., Wilkinson, D.T. & Corey, B.E. *Astrophys. J. Lett.* **232**, L139 (1979).
12. Hubble, E. *Proc. natn. Acad. Sci. U.S.A.* **22**, 621 (1936).
13. de Vaucouleurs, G. *IAU Symp.* No. 44, 353 (Reidel, Dordrecht, 1972).
14. *IAU Colloq.* No. 37, 119, 139, 141 (1977).
15. Soneira, R.M. *Astrophys. J. Lett.* **230**, L63 (1979).
16. Segal, J. *Mon. Not. R. astr. Soc.* **192**, 755 (1980).
17. Nicoll, J.F. & Segal, I.E. *Astr. Astrophys.* **82**, L3 (1980).

Parasitology in Venezuela

from Neil Brown and Win Gutteridge

THE study of parasitic diseases has for too long suffered from the geographical separation of pure and applied research efforts. Within the afflicted countries, research has concentrated on immediate clinical and epidemiological problems while more fundamental research has been confined to Europe and North America. As a consequence, both aspects of research have suffered, one through lack of modern know-how and the other through concentration on highly selective and convenient laboratory infections. A number of countries and international funding organizations have sought to end this separation and it is encouraging to report that, in Venezuela, a recent congress* demonstrated that fundamental research in parasitology is now being actively developed alongside clinical and epidemiological aspects.

Two protozoan diseases are of primary concern in Venezuela, Leishmaniasis and Chagas' disease. *Leishmania* are intracellular parasites of 'professional' phagocytes. Infection is initiated by a flagellated promastigote form inoculated by the sand-fly (*Phlebotomus*). Promastigotes actively secrete protein, and Hernandez and colleagues (Universidad Central de Venezuela, Caracas), in association with Greenblatt's group (Hadassah Medical School, Jerusalem), reported that as much as ten per cent of all protein synthesized is secreted. Although the secretory products have not been fully characterized, one, a glycoprotein with a high proportion of carbohydrate, has been shown by Hernandez' group to have the capacity to modify the membrane characteristics of macrophages.

Studies reported by Perez and associates (Instituto Venezolano de Investigaciones Cientificas, Caracas) have delineated a number of aspects of the host response to American cutaneous leishmaniasis. In syngeneic mice the type of lesion is determined mainly, but not exclusively, by the genetic composition of the host; parasite isolates also differ somewhat in their behaviour in one host genotype. Parasite destruction was shown by Perez to be mediated in an *in vivo* assay by peritoneal cells from immune animals and immune serum; non-adherent peritoneal cells were not protective. By contrast, in genetically susceptible BALB/c mice adherent cells were not protective and were found to have suppressor activity. Relating their laboratory studies to epidemiological observations, these workers were able to show that temporary protein deficiency ablates irreversibly the capacity of resistant C57BL/6 mice to resolve an active lesion. Other factors rather than simple killing of parasites by immune serum and activated macrophages must be involved in

---

*The 30th Convention of the Venezuelan Association for the Advancement of Science was held in Merida from 9 to 14 November 1980.

Neil Brown is in the Parasitology Division, National Institute for Medical Research, London and Win Gutteridge in the Biological Laboratory, University of Kent.

0028-0836/81/090746-02$01.00

Leishmania immunity, however, as even in C57BL/6 mice, lesions take 5–6 months to resolve, long after these immune responses appear. Practically nothing is known in detail about parasite antigenicity, particularly the antigenicity of the surface and secretory products of amastigotes, or the changes produced in the surface structure of the phagocytes they parasitize and eventually destroy. There is an urgent need to characterize the parasite side of the interaction and the role it plays in inducing parasite-specific and possible autoimmune protective and pathological reactions.

Little more is known of the biochemistry and molecular biology of *Leishmania* except in so far as it relates to energy metabolism, purine salvage and kinetoplast DNA. Ramirez and his colleagues (Universidad Central de Venezuela, Caracas) are now, however, beginning to probe the organization of the ribosomal RNA genes. They have estimated the molecular weight of the total genome to be 1.6×10^{11} (range for other protozoa 1–6×10^{11}) and that this includes about 350 copies of ribosomal DNA. An 8-megadalton fragment of total DNA, obtained by restriction endonuclease digestion with *Bam*HI and which hybridizes with ribosomal RNA, has been inserted into plasmid pBR322 and cloned in *Escherichia coli* x 1776. It should therefore be possible in future to study the ribosomal DNA of *Leishmania* in much more detail. In a complimentary investigation, Valbuena and his colleagues (Universidad Central de Venezuela, Caracas) have begun to study messenger RNA from the organism. They have developed a method for the isolation in high yield of polyribosomes which are active in a wheat germ-derived cell-free protein synthesizing system; messenger RNA is extracted from these.

Trypanosoma cruzi, the causative agent of Chagas' disease, is also an intracellular parasite, particularly of muscle tissue of the heart and alimentary tract where it produces severe disruption of organ function. As Torrealba (Universidad de Carabolo, Valencia) pointed out, it is a zoonosis which exists from the south of the US to the northern parts of Argentina and Chile. More than 150 species of mammal are carriers and serve as potential reservoirs of infection and more than 100 species of triatomid bug, the vector of transmission, are infected. In some areas of Latin America, infection rates in man are greater than fifty per cent.

A major problem for biochemists and immunologists studying parasites is obtaining adequate amounts of material. Avila (Pan America Center for Research, and Training in Leprosy and Disease of the Tropics, Caracas), who had already reported that epimastigote forms of *T. cruzi* can be grown *in vitro* in much simpler media than those currently in use, described the successful subculture of the organism in a medium containing salts, glucose, two growth factors and bovine liver catalase. The protein needed to be in its native configuration and was apparently not contaminated by purines, pyrimidines and growth factors. This observation suggests that the nutritional requirements of *T. cruzi* are markedly less exacting than those of other, non-endosymbiont-containing trypanosomatids. If confirmed, it will make life a lot easier for all *T. cruzi* investigators.

Piras and his colleagues (Centro Medico Docente La Trinidad, Caracas) made extensive use of probes such as lectins, microtubule disruptors and inhibitors of RNA and protein synthesis in their investigation of the mechanism of infection of fibroblastic cells by *T. cruzi.* The host cell membrane plays an active role in infection; in particular, surface glycoprotein(s) and an intact cytoskeleton must be present. Surface components of the parasite are also important; they may differ in each morphological form and some are possibly lacking in amastigotes isolated from cells since these are not infective. As low concentrations of actinomycin D and puromycin block infection it was suggested that both parasite and host cell membrane components need to interact and function if infection is to occur.

Considerable progress has been made in understanding the metabolism of *T. cruzi,* but few studies have been made of its regulation. Urbina and his colleagues (Universidad Central de Venezuela, Caracas) reported that *T. cruzi* contains a complete glycolytic sequence and tricarboxylic acid cycle but although glucose is metabolized, there is evidence for high rates of amino acid catabolism. Further investigations demonstrated that in *T. cruzi* there is little regulation of glycolysis by sugar phosphates or tricarboxylic acids but that amino catabolism is tightly regulated by energy charge.

Basic research in parasitology, using modern techniques, is developing rapidly in Venezuela. Most of the scientists running the programmes mentioned here have, as a result of national policy, spent some years in laboratories in Europe and North America. Clearly this policy is beginning to pay dividends. □

Weighing neutrinos: a small wait for a small weight?

from Frank E. Close

Do neutrinos have mass? The answer to this long-standing question may be answered within a few months if a proposal put forward by Alvaro De Rújula of CERN, Geneva is successful.

To provide the background to De Rújula's important idea it is worthwhile first recounting some of the history of neutrinos — mysterious particles produced in radioactive decays, which have no charge and (possibly) no mass, interacting so rarely with matter that they pass through the Earth as if it were empty space.

In β-radioactivity a neutron changes into a proton, and an electron is emitted to conserve electric charge:

$$n^0 \rightarrow p^+ + e^-$$

However, the energy and momentum of the proton and the electron do not equal that of the neutron. Rather than give up the principle of energy–momentum conservation, Pauli proposed, in 1931, that a third particle was produced in the process, and that this invisible particle (the 'neutrino') carried off the residual energy and momentum. The difference between the neutron mass and the mass of the proton plus electron would be the maximum mass that a neutrino could have and still be produced. Electrons were observed to be produced with energies up to about 1 MeV which showed that the neutrino was very light; indeed it could even have no mass at all. As no conclusive evidence for a mass has ever been found it is frequently referred to as 'massless'.

A quarter of a century passed before Cowan and Reines detected neutrinos being produced in the Savannah River Reactor in the US. These neutrinos were seen to induce the reverse reactions to those that had produced them in the first place. For their discovery they gained a case of champagne from Pauli — he had wagered that no one would ever detect so ghostly a particle!

Before turning to the ways that neutrino masses might be measured, the importance of massive neutrinos in cosmology is worth mentioning.

The density of neutrinos in the Universe should be similar to that of photons, that is at least 10^8 per nucleon. As a nucleon weighs about 10^9 eV, then neutrinos with masses even as small as 10–100 eV would provide the bulk of the Universe's mass! This would dramatically affect the expansion rate of the Universe and could

---

Frank E. Close is in the Theoretical Division of the Rutherford Laboratory, Didcot, UK.

0028-0836/81/090747-02$01.00

even slow it to the point where it stopped and then contracted to a big crunch — a replay of the big bang but in reverse. Cosmologists claim that the presently observed expansion rate probably limits the sum of the masses of all of the three known varieties of neutrinos to less than 100 eV, maybe less than 50 eV.

There have been some indirect hints that neutrinos have masses. The flux of neutrinos impinging on the Earth from the Sun's nuclear reactions (predominantly deuteron fusion $p + p \rightarrow de + \nu$) is a factor of 3 or 4 less than that expected by astrophysicists. This 'shortfall' can be understood if neutrinos have masses (see *Is the Sun still shining? Nature* **284**, 507; 1980).

The classic way to measure the neutrino mass has been through the study of the high-energy end of the spectrum of the electron produced along with the neutrino in the β-decay. At first sight this is an obvious approach — the heavier the neutrino, the lower will be the maximum available energy for the electron. The problem is that the neutron mass is not known to better than 40 eV and so one cannot resolve the neutrino mass to less than this. However, the *shape* of the high-energy tail of the electron energy spectrum depends on the mass of the accompanying neutrino (this is a kinematic property of phase space when three particles are produced as, for example, in $n \rightarrow p + e^- + \nu$).

Recently this technique has been employed in the USSR using tritium and studying the electron energy spectrum in

$$H^3 \rightarrow He^3 + e^- + \nu$$

This experiment claims the first non-zero mass result for ν — between 14 and 46 eV (V.A. Lubimov *et al. Phys. Lett.* **94B**, 266; 1980). Questions have been raised as to whether a zero mass is definitely ruled out.

In a widely attended seminar at CERN, de Rujula pointed out that there is another way that β-radioactivity produces neutrinos in a three-body final state which would be amenable to the 'shape dependence' test. Instead of producing an electron in $n \rightarrow p + e^- + \nu$ you capture an atomic electron. The nuclear capture of atomic electrons produces a continuous radiation spectrum

$$e^- + Z \rightarrow (Z-1) + \nu + \gamma$$

The photon plays the role that was previously played by the electron: measure the shape of the high-energy tail of the photon spectrum and you can infer the neutrino mass. The photon has no electrical charge and does not experience some of the scattering problems that beset charged electrons. Furthermore, detectors are available that can measure the photon energies with a precision of less than 1 eV!

If this was the whole story then one would have been doing such experiments long ago. The problem is that the event rate at the high-energy end of the spectrum tends to be so low that this approach is, at best, no improvement on the conventional technique. However, De Rújula noticed a feature of the radiation spectrum that could be exploited if a suitable nucleus existed and which would dramatically enhance the rate so allowing the experiment to be performed. Out of all the known isotopes it seems that Nature may have provided two that meet the crucial requirement! To explain this requirement we must start by detailing the nature of the radiation spectrum.

The first calculations of the intensity resulting from K-capture were made by Morrison and Schiff in 1940 (*Phys. Rev.* **58**, 24). While their calculations were in excellent agreement with the high-energy photon spectrum the observed spectra showed a considerably greater abundance of low-energy γ rays. Glauber and Martin (*Phys. Rev.* **104**, 158; 1956) showed that this enormous enhancement was due to radiative capture of electrons in L-states, the intensity peaking in the neighbourhood of the characteristic X-ray line. Thus for Fe^{55}, as an example, the peak is in the region of tens of keV while the low-intensity K capture continues all the way to the maximum available (230 keV).

If neutrinos have masses then the shape of the high-energy extreme of this distribution will be modified. Thus,although in principle we have a new way of determining the neutrino mass, the low event rate makes the method impractical. Can one find cases where the spectrum only extends for some tens of keV so that the L (or M) capture enhancement is at the high-energy end? In such circumstances the event rate will be enhanced in the very region that is relevant to the mass determination.

If you search through the atomic and nuclear tables very few examples are found. One possibility might be using

$$^{193}Pt_{78} \rightarrow {}^{193}Ir_{77} + 45\ \text{keV}$$

with a half life less than 500 years, but the most promising, studied in detail by De Rújula, seems to be

$$^{163}Ho_{67} \rightarrow {}^{163}Dy_{66} + \text{approximately zero}$$

with a half life which has not been well determined. In this latter case the high-energy end of the photon spectrum will come entirely from L,M . . . capture and have high intensity, so yielding excellent event rates.

CERN is particularly well suited to performing such an experiment.

First of all it possesses an isotope on-line separator (ISOLDE). The 600 MeV proton beam from the CERN synchocyclatron is focused on a target ion source. A beam of radioactive ions is formed by a 60 kV acceleration stage and is mass analysed by a magnet. Individual mass isotopes can then be selected by electrostatic deflection, in particular the desired Holmium isotope crucial to the electron capture experiment.

The energy spectrum of the emitted photons can be studied with an accuracy of order 1 eV, and the shape of the high-energy tail thereby precisely determined.

It is known that (at least) three varieties of neutrino exist in Nature — ν_e, ν_μ, ν_τ of which the ν_e variety is produced in the above experiments. However, each of these is, in principle, a linear superposition of mass eigenstates ν_1 ν_2 ν_3. As a result the ν_e is produced with various (unknown) probabilities to have mass m_1 m_2 or m_3. All three cases may affect the energy spectrum to produce a series of kinks. From the positions of the kinks it could be possible to determine the masses and relative probabilities for the various states.

The accumulation of isotopes is now in progress and the optimists claim that the experiment can be performed very quickly. It is possible that we have only a small time to wait before the neutrino masses are measured — unless some unforseen experimental problems arise. ☐

Woodchuck, squirrel and duck hepatitis viruses

from Arie J. Zuckerman

THE close phylogenetic relationship between human hepatitis B virus and a virus isolated from the eastern woodchuck (*Marmota monax*) has been clearly established by Cummings and his associates (*Proc. natn. Acad. Sci. U.S.A.* **77**, 1842; 1980). They have cloned the DNA of human hepatitis B virus and the DNA of woodchuck hepatitis virus in the vector γgtWES and subcloned into the kanamycin-resistant plasmid pAO1. Comparison of the recombinant DNAs with authentic virus DNAs by specific hybridization, size and restriction enzyme analysis indicated that the recombinants contained the complete genome of each virus. The nucleic acid homology between the two viral DNAs was confirmed with the cloned DNAs. Infection with either virus results in the accumulation in blood of

---

Arie J. Zuckerman is Professor of Microbiology and Director of the WHO Collaborating Centre for Reference and Research on Viral Hepatitis at the London School of Hygiene and Tropical Medicine.

0028-0836/81/090748-02$01.00

large amounts of excess virus coat protein in the form of spherical and tubular particles measuring 20–25 nm in diameter, and both viruses have 40–50 nm double-shelled or solid particles with a nucleocapsid or core containing double-stranded circular DNA with a gap and a virus DNA polymerase, and show a high degree of antigenic cross-reactivity between the cores of the two viruses, and a small region of 100–150 base pairs of nucleic acid homology as measured by liquid hybridization. The analogy between the two viruses is even closer when judged by their adaptation in their respective hosts causing persistent infection and close association with chronic liver disease and primary liver cancer.

Yet another virus related to human hepatitis B in Beechey ground squirrels (*Spermophilus beecheyi*) has been found in one region in northern California. Marion and her colleagues (*Proc. natn. Acad. Sci. U.S.A.* **177**, 2941; 1980) identified a virus in these squirrels sharing many of the unique characteristics of the human virus. Common features include virus morphology, size and structure of the viral DNA, a virion DNA polymerase which repairs a single-stranded region in the double-stranded circular genome, cross-reacting surface viral antigens, antigen–antibody systems similar to hepatitis B *e* antigen and the core antigen, and persistent infection with viral antigen present continuously in the blood. Because the ground squirrels were only bled and then released, it has not been possible to observe a disease accompanying this virus in the ground squirrel.

There is also some preliminary evidence suggesting further members of this unusual class of viruses in black-tailed prairie dogs (*Cynons ludoviccianus*) and in ducks in a province in the People's Republic of China. Virological studies, however, have not yet been completed.

As the human hepatitis B virus has not been grown in tissue culture, these phylogenetically related animal viruses may provide valuable models for the study of viral replication and the pathogenesis of chronic liver disease and primary liver cancer. □

Hominoid habitats of the Miocene

from Peter Andrews

NEW STUDIES on the palaeoecology of fossil hominoids make it increasingly likely that fossil hominoids of the early Miocene, 18 to 20 million years ago, lived in quite different habitats from those of the middle Miocene, 12 to 15 million years ago, and suggest that the way of life of the ancestors of the great apes and man will have to be reconsidered. The early Miocene hominoid species are generally found associated with fossils indicating tropical forest conditions in contrast with the middle Miocene hominoids which are associated with temperate to tropical woodland. This ecological shift is accompanied by changes in tooth morphology — in particular the development of thick enamel — and in many other characters of the skull and postcrania. The early Miocene hominoids retain mainly primitive characters while the middle Miocene species have developed many more of the derived characters shared with the living great apes and man.

Hominoids in the middle Miocene were widely distributed in Europe and Asia along a broad belt stretching from southern Europe across to south-western China. Much of this region was covered with broad-leaved sclerophyllous woodland related to a vegetation type found today only in the Canary Islands (Axelrod *Ann. Miss. Bot. Gard.* **62**, 280; 1975). The vegetation type was intermediate between the tropical and subtropical forests to the south and the moist temperate woodlands with mixed conifers and broad-leaved trees to the north. It has not been clear to which part of this spectrum the middle Miocene hominoids were adapted, because direct associations betwen hominoid species and fossil floras have not, until recently, been available. Interpretations of past habitats by analysis of the animal faunas are limited, and at the *Ramapithecus* and *Sivapithecus* localities they do no more than indicate some type of woodland. However, a recently described flora from China that comes from the same beds as a mammalian fauna containing these hominoids provides the first direct evidence of the vegetation type associated with the middle Miocene thick enamelled apes.

The flora frum Lufeng, in south-western China, shows a succession up the fossiliferous sequence from a warm humid vegetation type to cool temperate types (Sun Xiang-jun and Wu Yu-shu *Vert. Palas.* **18**, 255; 1980). The middle part of the sequence, which is where the hominoids occur, is a lignite containing pollen of both aquatic and land plants. The lignitic nature of the deposits and the presence of aquatic plants suggest swamp conditions but by far the most abundant plant forms are temperate trees and shrubs represented by 44 taxa, with 8 taxa of ferns, 11 of herbs, and 3 aquatic forms. The trees and shrubs consist of both evergreen and deciduous taxa, and the composition of the flora, although containing some sclerophyls, is closest to that of the temperate woodlands described by Axelrod as occurring to the north of the sclerophyllous belt during the middle Miocene. Woodlands of this type occur today in southern Europe to the north of the Mediterranean and Black seas. It is an unusual vegetation type in which to find primates, now largely tropical and subtropical in distribution and suggests that enamelled hominoids must have been able to survive in conditions beyond the range of most living primates.

Thick enamelled hominoids are known during the middle Miocene from Africa as well as from Europe and Asia. Faunas from Fort Ternan and Maboko Island, Kenya, have been analysed in another recent paper (Nesbit Evans *et al. J. Hum. Evol.* 10 (1), 1981), and demonstrate the presence of tropical woodland. The faunas have been analysed in two ways — by assessment of indicator species in the faunas, either singly or by combinations of species to produce habitat spectra, and by analysis of the community structure of the mammalian faunas through quantification of the diversity of ecological adaptations present in the faunas. The results of the two methods agree with each other, and with other evidence from the lower vertebrates and invertebrates, and assuming that the association between the hominoid species and the rest of the fauna is valid it can be inferred that the hominoids (*Ramapithecus* and/or *Sivapithecus*) also occupied woodland habitats during the middle Miocene in Africa.

The evidence for several early Miocene sites in Kenya is also reviewed in the same paper following the same methods. These confirm earlier results that at certain of the levels at Songhor, Rusinga Island and Koru, where the faunas are associated with hominoid species of the genera *Proconsul*, *Limnopithecus* and *Dendropithecus*, tropical forest habitat occurred.

The new data are of importance in considering the evolutionary emergence of the living great apes and man for they imply that their ancestors passed through an adaptive phase very different from their present way of life. *Sivapithecus* and *Ramapithecus* were the size of orang-utans and baboons respectively and animals of this size living in woodland habitats, whether tropical or temperate, must have been at least partly terrestial because of discontinuities in the woodland canopy. Again, because woodland habitat shows greater seasonal variation than forest the supply of fruit would have been insufficient to have played more than a minor role in the diet. In conclusion, many previous studies that have claimed an arboreal and frugivorous ancestry for the living great apes and man will need to be reconsidered, for the evidence from Middle Miocene suggests that hominoids of this period were omnivorous and partially terrestrial. □

Peter Andrews is in the British Museum (Natural History), London.

0028-0836/81/090749-01$01.00

Variability of the Earth's rotation

from Frank D. Stacey

LONG BEFORE the advent of atomic clocks, it was well known that the Earth's rotation was an imperfect time-keeper. Precise time-keeping has always been one of the roles of astronomy and the stars, planets and satellites provide us with several independent or semi-independent clocks. Rather than clarifying the problem, modern measurements and recent re-analyses of earlier data have found evidence of a multiplicity of superimposed mechanisms. The difficulty is partly, of course, that some of the mechanisms involve time constants that are longer than the period of observation and so, regardless of the precision of observations, a long record is required to sort them out.

The diversity of disciplines that contribute to the study of the rotation of the Earth makes it an interesting theme for a wide-ranging discussion and, at a recent meeting*, astronomy, meteorology, oceanography, palaeontology, cosmology and the physics of the Earth's deep interior were all well represented. New data and ideas were forthcoming, as well as some thoughtful questioning of arguments supposed to be reasonably secure.

It now appears that at least some of the palaeontological data upon which we have relied for direct evidence of the relative lengths of the day, month and year in remote geological periods are very unsure. Certainly growth increments on the shells of marine organisms are well observed, but do they really indicate diurnal, monthly and annual periods, with all time increments faithfully recorded? Some of these questions have recently been brought to the fore by a controversial study of nautiloid growth rhythms (P. G. K. Kahn and S. M. Pompea *Nature* **275**, 606; 1978; see also *Nature* **279**, 452; 1979). W. W. Hughes (Andrew University, Michigan) has been examining a large number of nautiloid specimens from the British Museum (Natural History) collection. At least some organisms record the semi-diurnal lunar tide rather than the day and those observed under controlled conditions give no observable growth increments during the winter months. While we are not ready to give up hope of definitive data from growth rhythms, clearly we need much more work on control specimens and interpretation of ancient samples in the light of it before estimates of rotation rates in earlier epochs can be regarded as secure.

The gradual slowing of the Earth's rotation by tidal friction can be measured in several ways although the obvious and direct measurements of the apparent motion of distant stars with precise clocks (atomic or astronomical) is confused by shorter term fluctuations due to exchange of angular momentum within the Earth. It is agreed that tidal friction in the Earth is dominated by the dissipation of tidal energy in the sea. To extrapolate to the remote past a linear relationship, in the sense that the lunar and solar tides act independently, has generally been assumed, with dissipation proportional to the square of tidal amplitude. For most purposes this may not be too bad an approximation, in spite of the evidence that tides are turbulent and therefore at least to some extent non-linear. However, when fine details of the angular momentum exchange with the lunar and solar orbits are sought, this assumption becomes unsatisfactory. For some years T. C. Van Flandern, at the Naval Observatory in Washington, has sought, with progressively increasing precision, evidence for a time variation of the newtonian gravitational constant, G, in terms of an atomic clock determination of the lunar orbital acceleration. It is important to know whether such an observation may be confused by non-linearity of the tide because the length of the record of atomic-timed observations of the Moon appears to be approaching the point of giving a definite result. According to V. Canuto (NASA Goddard Space Flight Center, *Nature* in the press) and G. M. Blake (University of Sheffield), the bias of presently available data favours an increase in G with time rather than the decrease postulated by P. A. M. Dirac.

Evidence of short-term variations in rotation is improving rapidly. A new analysis by L. V. Morrison (Royal Greenwich Observatory) and F. R. Stephenson (Liverpool University) of the timing of occultations, with some eclipse data back to 1620, gave a much more convincing length-of-day curve than any hitherto for the period from about 1750; the data before that time were of doubtful value. There are no discontinuities like those that appeared in earlier analyses, but the occurrence of changes amounting to a few parts in 10^8 over 5 or 10 years is strikingly confirmed. Irregular motions in the Earth's core are clearly implicated although the mechanism is a subject of controversy. A new examination from French observatory going back to 1700 of the correlation between rotational changes and the geomagnetic westward drift by J. L. Le Mouël and V. Courtillot (IPG, University of Paris) showed that magnetic changes preceded rotational changes by two or three years, contrary to some questionable analyses reported earlier. This indicates a semi-independent electromagnetic coupling of the Earth's solid mantle to the outer region of the core (which carries the westward drifting features of the field, as originally proposed by E. C. Bullard), and, more strongly, to the deeper body of the core. There are clearly difficulties with this hypothesis, although it does provide a quantitative explanation of the rotation observations.

The time scale for core-induced changes in rotation rate is several years and a coupling time-constant of this magnitude is compatible with the electromagnetic torques exerted by the core-generated field on a lower mantle of conductivity $100\Omega^{-1}$ m^{-1} or so. However, the observations permit the supposition that core–mantle coupling is very tight indeed and that the apparent time constant is a feature of the core motions themselves. This supposition is necessary if the core has any role in exciting or damping the Chandler wobble. The conventional wisdom follows the Rochester–Smylie analysis of wobble coupling, which demonstrates that for the same field and mantle conductivity configuration the coupling time constant for the wobble is about 100 times as long as for the length-of-day changes. This is probably the strongest single reason for believing that tectonic displacements, including earthquakes, within the mantle excite the wobble, but S. K. Runcorn (University of Newcastle upon Tyne) keeps alive the argument that the core may be responsible for the wobble. The necessary mantle layer with a conductivity nearer to $10^4\Omega^{-1}m^{-1}$ is not well supported by observations, but such a layer does have proponents for other reasons, so this particular debate will continue for a while yet.

The smaller, but much more rapid, fluctuations in rotation must be attributed to atmospheric motions. A comprehensive new global analysis of the atmospheric angular momentum, reported by R. Hide, (Meteorological Office, Bracknell; see also *Nature* **286**, 114; 1980; **286**, 104; 1980) confirmed that major changes can occur quite rapidly, reacting on the Earth in a matter of days. The coupling is obscure but must certainly be there if the angular momentum exchange is demonstrated. A comparable analysis of the perpendicular components of atmospheric angular momentum, coupling to which would excite the wobble instead of changes in rotation rate, is promised. This reopens another of those debates that has been closed and reopened several times in the past 25 years.

What are the outstanding problems for the immediate future? If we compile a complete list, practically every problem that has ever been tackled is still on it. A personal list, restricted to four items in order of significance is (i) Adequacy of newtonian mechanics to explain the lunar orbit, (ii) Periodicities in the growth patterns of organisms exposed to a tidal cycle, (iii) Atmospheric excitation of the wobble, (iv) Lower mantle conductivity and core coupling. □

Frank D. Stacey is Professor of Applied Physics at the University of Queensland, Australia.

---

*Organized by Professor S.K. Runcorn for The Royal Astronomical Society on 9th January 1981.

At NEN, service is not just a policy: It is an attitude.

Highest specific activity CATECHOLAMINES

Epinephrine, *levo*-[*N-methyl*, *ring*-2,5,6-^{3}H]-
95-145Ci/mmol
0.2N Acetic acid:ethanol, 9:1, under nitrogen, foil-wrapped vial, dry ice
NET-696 250μCi 1mCi

Norepineprine, *levo*-[*ring*-2,5,6-^{3}H]-
40-60Ci/mmol
0.2N Acetic acid:ethanol, 9:1, under nitrogen, foil-wrapped vial, dry ice
NET-678 250μCi 1mCi 5mCi

Dihydroxyphenylethylamine hydrochloride, 3,4-[*ring*-2,5,6-^{3}H(N)]-
40-60Ci/mmol
Ethanol:0.2N Acetic acid, 1:9, under argon, foil-wrapped vial, dry ice
NET-673 250μCi 1mCi

Circle No.19 on Reader Enquiry Card.

NEWEST LIGANDS

Call or write for our complete list of over 100.

Adenosine receptor agonist
Cyclohexyladenosine, N^{6}-[*adenine*-2,8-^{3}H]-
Adenosine receptor antagonist
Diethyl-8-phenylxanthine, 1,3-[*phenyl*-4-^{3}H]-
Serotonin-2 antagonist
Mianserin hydrochloride, [*N-methyl*-^{3}H]-
GABA agonist
Piperidine-4-sulfonic acid, [*ring*-^{3}H]-
Irreversible cholinergic antagonist
Propylbenzilylcholine mustard hydrochloride, [*propyl*-2,3-^{3}H]-
Histamine-2 antagonist
Tiotidine, [*methyl*-^{3}H]- **(ICI 125,211)**
Potential barbiturate receptor ligand
(1,3-Dimethylbutyl)-5-ethylbarbituric acid, (−)-5, [*butyl*-2,3,4-^{3}H (N)]-

Circle No. 42 on Reader Enquiry Card.

Custom design and synthesis of PHOTOAFFINITY PROBES

Labeled in alanine-1 position to prevent possible loss of label by hydrolysis

Azido-2-nitrophenyl-β-alanine, *N*-4-[*alanine*-1-^{14}C]-
40-60mCi/mmol
Ethanol:water, 7:3, foil-wrapped vial, dry ice
NEC-769 50μCi 250μCi

Circle No. 43 on Reader Enquiry Card.

TRH [^{125}I]

Carrier free. Tested for binding and displacement in a specific RIA. Standard curve supplied with shipment.

Thyrotropin releasing hormone, [^{125}I]- (monoiodinate
pGlu-**His**-Pro-NH$_2$
~2000Ci/mmol carrier free
Propanol:water, 1:1
NEX-153 10μCi 2x10μCi 50μCi 2x50μCi
Prepared fresh for stock 2nd Monday each month

Circle No. 44 on Reader Enquiry Card.

Substance P (8-L-tyrosine) [^{125}I]

High specific activity >500μCi/μg
Stable four weeks from production date
Each lot tested for binding and displacement in a specific RIA

Substance P (8-L-tyrosine), [^{125}I]-
>500μCi/μg
Lyophilized from sodium carbonate buffer, pH 9.9, containing bovine serum albumin
NEX-152 10μCi 2 x 10μCi 50μCi 2 x 50μCi
Prepared fresh for stock third Monday each month

Circle No. 45 on Reader Enquiry Card.

Not for use in humans or clinical diagnosis

NEN New England Nuclear®
549 Albany Street, Boston, Mass. 02118
Call toll-free: 800-225-1572
(In Massachusetts and International: 617-482-9595)

NEN Chemicals GmbH: D-6072 Dreieich, W. Germany
Postfach 401240, Tel. (06103) 85034, Telex 4-17993 NEN D
NEN Canada: 2453 46th Avenue, Lachine, Que. H8T 3C9
Tel. 514-636-4971, Telex 05-821808

Isn't it Time

to have a subscription to Cell? Now in its seventh year, Cell is the prestige journal of modern biology, publishing seminal research papers on topics ranging from molecular biology to developmental biology, from cell biology to virology and genetics. Many important discoveries were first reported in Cell; current examples include such diverse reports as the first cloning of the mouse and the discovery of a splicing protein coded by an intron.

Minireviews, reviews and book reviews also are a prominent part of Cell, keeping readers abreast of research and other events in all areas. With the pace of research so fast, this part of the journal is essential reading for all researchers.

Cell is so much in demand that copies may be hard or impossible to find in the library. (Some libraries keep current issues under lock and key.) Personal subscriptions offer immediate access to the latest research and availability for future reference. *And personal subscriptions are available at less than half the regular rate.* The order form below can be used to obtain one starting with the January issue. To keep abreast of the demand to publish in Cell, there will be 15 issues in 1981.

Cell Personal Order Form

To: The Journals Dept.
MIT Press
28 Carleton St
Cambridge, MA 02142 USA

I wish to obtain a personal subscription to Cell for 1981. I enclose a personal check for $90 ($100 outside the US) and declare that the subscription is for my own use only.

Signed ______________________________

Name ______________________________

Address ______________________________

____

____

N5

Circle No.46 on Reader Enquiry Card.

Nature Vol. 289 26 February 1981

REVIEW ARTICLE

Attenuation in the control of expression of bacterial operons

Charles Yanofsky

Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

Bacterial operons concerned with the biosynthesis of amino acids are often controlled by a process of attenuation. The translation product of the initial segment of the transcript of each operon is a peptide rich in the amino acid that the particular operon controls. If the amino acid is in short supply translation is stalled at the relevant codons of the transcript long enough for the succeeding segment of the transcript to form secondary structures that allow the transcribing RNA polymerase molecule to proceed through a site that otherwise dictates termination of transcription. This site is the attenuator; the process is attenuation.

IT has long been known that RNA polymerase, with and without accessory factors, recognizes specific nucleotide sequences which signal the initiation and termination of transcription, and that this recognition is exploited in various ways in the regulation of gene expression. But during the past 10 years it has become evident that transcription termination sites can be located within as well as at the end of an operon. These internal termination sites are used in the regulation of transcription and may be used to sense features and components of metabolism that cannot be used to control initiation of transcription. This mechanism of regulation, called attenuation, can be defined as the regulation of gene expression by selective reduction of the transcription of distal portions of an operon. Regulatory studies with bacteriophage λ (refs 1–5) revealed the first case of attenuation. Two operons of this phage contain transcription termination sites which are used to control transcription of the structural genes located immediately beyond. Soon after these studies with λ a different mechanism of attenuation was discovered in the tryptophan[6] (*trp*) and histidine[7] (*his*) operons of bacteria and subsequently in several other operons of the enzymes of amino acid biosynthesis. In this review I will discuss the mechanisms of attenuation, concentrating on the tryptophan operon of *Escherichia coli*.

During the early history of studies on gene expression in amino acid biosynthetic operons, thoughts on possible regulatory mechanisms were dominated by the conceptual contributions of Jacob and Monod. The focus of investigation was therefore on the detection of a repression system and on the relative roles of amino acid and charged transfer RNAs as activators of a presumed repressor protein[8,9]. Nevertheless, Ames and his co-workers concluded that transcription of the *his* operon of *Salmonella typhimurium* was probably not governed by a repressor protein and that tRNA^{His} rather than histidine was the signal molecule[10,11]. Martin, Ames and Hartman[12], in some early speculations, proposed that it was translation of the transcript of the initial segment of the *his* operon that regulated transcription of the remainder of the operon. During the same period investigations on *trp* operon expression in *E. coli* followed a different course because there was convincing evidence that a tryptophan-activated repressor protein regulated transcription of the operon[13–19]. However, several observations were inconsistent with the view that repression was the sole regulatory mechanism of this operon. Tryptophanyl-tRNA synthetase mutants were shown to have regulatory anomalies[20–23]; addition of tryptophan to these strains did not fully depress *trp* operon expression. Moreover, Imamoto observed that the addition of tryptophan to tryptophan-starved cells not only shut down the initiation of transcription but also inhibited transcription in progress on the initial segment of the operon[24,25]. Finally, mutants lacking a functional repressor could still respond to tryptophan starvation by increasing their rate of synthesis of *trp* messenger RNA[26,27].

Definitions of terms

Leader region—The segment of an operon between the transcription start site and the structural gene(s).

Attenuator—a transcription termination site within an operon.

Transcription pause site—a region of DNA at which RNA polymerase pauses in the course of transcription.

Terminated leader transcript—The transcript terminated at the attenuator.

Termination structure—The segment of the leader transcript that is thought to be recognized as a termination signal by RNA polymerase.

Leader peptide—The short peptide encoded in the terminated and readthrough transcripts of the leader region.

These findings suggested that repression was not the sole means of regulating the *trp* operon of *E. coli*. Ultimately, two observations forced consideration of alternative regulatory schemes. Among a set of *trp* deletion mutants, in which both deletion termini were within the transcribed region of the operon, some had an unexpected sixfold increase in expression of the remaining genes of the operon[6,28]. Since repressor control was not affected by these deletions, what must have been deleted was a site distinct from the operator but which was also used to regulate transcription of the operon. The second observation was that within the mRNA sequences produced *in vivo* from the 5′ end of the *trp* operon, oligoribonucleotides corresponding to the first 140 base pairs of the operon were several times more abundant than those derived from more distal segments[29,30]. These findings suggested that there was a site of transcription discontinuity, possibly a stop site, in the initial segment of the operon[6].

When transcription of the initial segment of the *trp* operon was studied *in vivo* and *in vitro*[28,30–34] it became apparent that a transcription termination site was located before the structural genes of the operon. Moreover, starving bacteria of tryptophan reduced termination of transcription at this site[27,35]. Kasai provided convincing evidence of an equivalent site of regulation of transcription of the *his* operon of *S. typhimurium*[7]. He called regulation at this site 'attenuation', the term now generally used to describe regulation by transcription termination.

0028-0836/81/090751—08$01.00

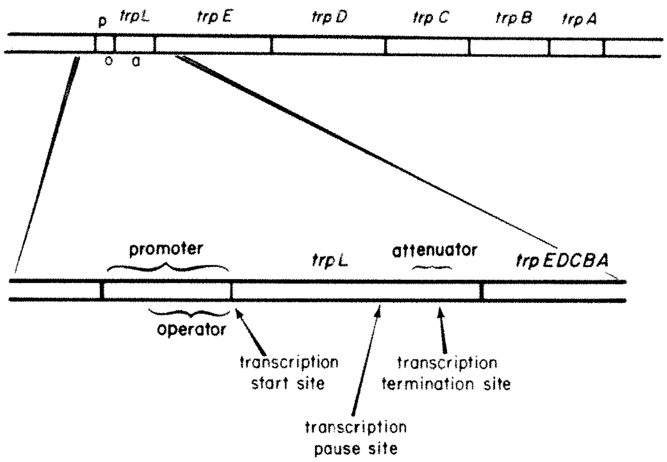


Fig. 1 The regulatory and structural gene regions of the *trp* operon of *E. coli*. Transcription initiation is controlled at a promoter–operator. Transcription termination is regulated at an attenuator, in the transcribed 162-base pair leader region, *trpL*. All RNA polymerase molecules transcribing the operon pause at the transcription pause site before proceeding further.

Features of the tryptophan operon and its regulatory region

The *trp* operon of *E. coli* consists of a transcription regulatory region and five structural genes encoding the polypeptides that catalyse the terminal sequence of reactions in tryptophan formation[36,37] (Fig. 1). The operon is approximately 7,000 nucleotide pairs long, and each cistron of the polycistronic transcript of the operon seems to be translated equally. The initiation of transcription is regulated at an operator site located within the promoter region of the operon[15–20,38,39]. A tryptophan-activated repressor protein, the product of the unlinked regulatory gene *trpR*, regulates initiation by controlling the access of RNA polymerase to the promoter[39–41]. The transcription termination site or attenuator (see below for definitions of terms) is located within the 162-nucleotide pair transcribed leader region, *trpL*, preceding the first major structural gene. There is also a transcription pause site within *trpL* (M. Winkler and C. Yanofsky, in preparation). A possible role for this site will be discussed below.

Repression rather than attenuation exerts the greater quantitative effect *in vivo* on transcription of the structural genes of the operon; repression can reduce transcription up to 70-fold[42] while attenuation can only reduce it 8–10-fold[6,35]. Because tryptophan starvation increases expression by relieving both repression and attenuation, transcription of the *trp* operon of *E. coli* can be regulated over a range of about 600-fold. For a more thorough discussion of all aspects of regulation of tryptophan biosynthesis in bacteria the reader is referred to the excellent reviews of Crawford and Stauffer[43] and Platt[37].

Attenuation

In addition to locating the attenuator in the distal segment of *trpL*[6,28], mutants have also provided clues to the mechanism of attenuation. Mutations altering tryptophanyl-tRNA synthetase, tRNA^{Trp}, or a tRNA^{Trp} modifying enzyme and thereby reducing the efficiency of translation of Trp codons, were also found to decrease the termination of transcription at the *trp* attenuator[44–47]. This suggested that attenuation involved the translation of Trp codons rather than simply the recognition of changes in the tryptophan concentration. Attenuation in the *trp* operon is also affected by mutations of functionally important sequences in the leader region. Finally, mutations affecting RNA polymerase and the transcription termination protein rho also affect termination of transcription at the *trp* attenuator[47,48]. The role of rho, if it has one, in transcription termination at the *trp* attenuator is unclear.

The leader transcript

Both *in vivo* and *in vitro* transcription studies established that transcription can be terminated in the leader region of the *trp* operon of *E. coli*[28,32]. Termination occurs at approximately the same position *in vivo* and *in vitro* to give a leader transcript approximately 140 nucleotides in length[31]. The DNA region within which RNA synthesis stops is A+T rich and is immediately preceded by a G+C-rich region that exhibits dyad symmetry. This arrangement is typical of many prokaryotic transcription termination sites[49] and both regions are implicated in the transcription termination event. The question is how, at the *trp* operon attenuator, can tryptophan deficiency be sensed and communicated to the transcribing RNA polymerase molecule, allowing it to transcribe beyond the attenuator into the structural genes of the operon?

Ribosome binding experiments carried out with the 140 nucleotide *trp*-leader transcript yielded the unexpected result that ribosomes protect a 20-base segment of the early portion of the transcript from nuclease attack[50] (Fig. 2). A potential AUG start codon is located in the centre of the protected region. Were translation initiated at this start codon and terminated at the next stop codon (UGA at nucleotides 69–71), a 14-residue peptide would be synthesized with adjacent tryptophan residues near its distal end (Fig. 2). The presence of tandem tryptophan residues in this predicted 14-residue peptide was tantalizing because tryptophan is a rare amino acid in the proteins of *E. coli*, generally being used only once in every 100 amino acid residues. (The amino acid constitution of the predicted leader peptides now known to be specified by other biosynthetic operons (Fig. 3) turned out to be even more striking—see below.) The presence of the tandem tryptophan residues in this small peptide encoded in the leader region suggested a possible role for tRNA^{Trp} in the regulation of transcription termination at the *trp* operon attenuator[33]. It was hypothesized that in cells starved of tryptophan a ribosome translating the leader transcript would stall at either of the two Trp codons. Only with sufficient tryptophan would the ribosome complete the entire leader peptide. The final position of the translating ribosome on the transcript could then be the distinctive feature communicated to the transcribing polymerase in the regulation of termination at the attenuator. It was reasonable to consider this hypothesis since in prokaryotes transcription and translation proceed in concert. These considerations raised two crucial questions. Is the potential AUG start codon within the ribosome-protected region (the ribosome binding site) actually used as a site of translation initiation, and, is the existence of adjacent Trp codons in the transcript coincidental?

The *trp* leader ribosome binding site was shown to be an efficient site for the initiation of translation by fusing it to a structural gene or a portion thereof and demonstrating the synthesis of fused polypeptides[51,52]. Since the leader ribosome binding site is used, we must presume that our inability to detect

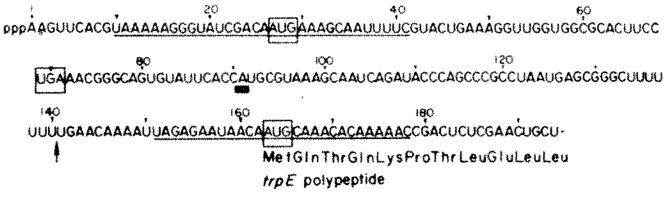

leader peptide
MetLys Ala Ile PheValLeuLys Gly Trp Trp ArgThr Ser
1
20
40
60
pppAAGUUCACGUAAAAAGGGUAUCGACAAUGAAAGCAAUUUUCGUACUGAAAGGUUGGUGGCGCACUUCC
80
100
120
UGAAACGGGCAGUGUAUUCACCAUGCGUAAAGCAAUCAGAUACCCAGCCCGCCUAAUGAGCGGGCUUUU
140
160
180
UUUUGAACAAAAUUAGAGAAUAACAAUGCAAACACAAAAACCGACUCUCGAACUGCU-
MetGlnThrGlnLysProThrLeuGluLeuLeu
trpE polypeptide

Fig. 2 The nucleotide sequence of the 5′ end of *trp* messenger RNA. The non-terminated transcript is presented. When transcription is terminated at the attenuator, a 140-nucleotide transcript is produced. Its 3′ terminus is marked by an arrow. The 3′ terminus of the pause transcript, at nucleotide 90, is underlined by a bar. The two AUG-centred ribosome binding sites in this transcript segment are underlined. The boxed AUGs are where translation starts and the boxed UGA where it stops. The predicted amino acid sequence of the *trp* leader peptide is shown.

phe, *his*, *leu*, *thr* and *ilv* leader peptides

pheA:	Met-Lys-His-Ile-Pro-*PHE-PHE-PHE*-Ala-*PHE-PHE-PHE*-Thr-*PHE*-Pro
his:	Met-Thr-Arg-Val-Gln-Phe-Lys-*HIS-HIS-HIS-HIS-HIS-HIS-HIS*-Pro-Asp
leu:	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-*LEU-LEU-LEU-LEU*-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His
thr:	Met-Lys-Arg-*ILE*-Ser-*THR-THR-ILE-THR-THR-THR-ILE-THR-ILE-THR-THR*-Gly-Asn-Gly-Ala-Cly
ilv:	Met-Thr-Ala-*LEU-LEU*-Arg-*VAL-ILE*-Ser-*LEU-VAL-VAL-ILE*-Ser-*VAL-VAL-VAL-ILE-ILE-ILE*-Pro-Pro-Cys-Gly-Ala-Ala-Leu-Gly-Arg-Gly-Lys-Ala

Fig. 3 The predicted amino acid sequences of the leader peptides of the *pheA*, *his*, *leu*, *thr* and *ilv* operons of *E. coli* or *S. typhimurium*. Amino acids which regulate the respective operons are in italicized capitals and are underlined. For references see text.

the leader peptide is due to its instability. That the leader peptide as such does not have a regulatory role is indicated by complementation analyses[53].

The significance of the tandem Trp codons in the *trp* leader transcript of *E. coli* is supported by their presence also in the *trp* leader regions of *Shigella dysenteriae*, *Salmonella typhimurium*, *Serratia marcescens*[34,54,55] and *Klebsiella aerogenes* (M. Blumenberg and C. Yanofsky, unpublished). The predicted amino acid residues of these *trp* leader peptides are also conserved in the Arg-Thr-Ser sequence that follows the adjacent Trp residues but not in the sequence that precedes it. More importantly, the predicted leader peptides of other amino acid biosynthetic operons that are regulated by attenuation also are rich in the amino acid which is the end product of the appropriate pathway (Fig. 3). Hence, in a bacterial cell starved of any of these amino acids, the ribosome translating the leader transcript of the respective operon would stall over the appropriate codon rather than move to the leader transcript stop codon. Figure 3 reveals that the number as well as the location of the distinctive amino acids varies, presumably reflecting quantitative differences in regulation of the different operons.

If ribosome stalling at the Trp codons of the transcript of the leader region were sufficient to prevent the termination of transcription at the *trp* attenuator, then would stalling anywhere in the coding region of the transcript have the same effect? Starving *E. coli* mutants for each of seven amino acids encoded in the *trp* leader transcript (Met, Val, Leu, Gly, Trp, Arg and Thr) and two that are not (His and Pro)[53] revealed that the starvation response was relatively specific. Only cells depleted of Trp or Arg exhibited reduced termination of transcription at the *trp* attenuator. Comparable starvation experiments with the *trp* operon of *S. marcescens* (I. Stroynowski and C. Yanofsky, unpublished) demonstrated that starvation for His, the second residue before Trp in the leader peptide, as well as for Trp or Arg, reduced the termination of transcription at the *trp* attenuator. These observations suggested that ribosome stalling in the vicinity of the Trp codons is sufficient to exert a regulatory effect.

This conclusion led to the question of how amino acid starvation and presumed ribosome stalling are communicated to the transcribing polymerase. Hoping to answer this question we scrutinized the transcript of the leader region, searching for potential secondary structures that might relate to the transcription termination event. This search was prompted by the suggestion of Marty Rosenberg, on the basis of sequence analyses with a terminated transcript of bacteriophage λ[56], that base-paired segments in RNA may serve as a signal in the transcription termination event. Limited RNase T_1 digestion of the 140-nucleotide long *trp* leader transcript revealed extensive secondary structure in the transcript; many phosphodiester bonds in the latter half of the transcript were relatively resistant to nuclease attack[33]. Also, certain fragments in partial digests of the transcript were observed to migrate together during electrophoresis in non-denaturing gels[57]. Presumably they did so because they were base-paired. Analyses of the nuclease-resistant RNA fragments, as well as computer matching of complementary sequences, indicated that the transcript contains the two hairpin loops shown in Fig. 4 (left). For simplicity, we designate the four strands participating in these structures 1–4 (Fig. 4). We observed that strands 1 and 2 were paired and that strand 3 remained covalently attached to strand 4. The calculated free energies of these base-paired structures are given in Fig. 4. Note that strand 1 contains the adjacent Trp codons and that structure 3:4 resembles other hydrogen-bonded hairpin structures that immediately precede the 3′ ends of terminated transcripts. Inspection of the sequences of the paired regions revealed an interesting feature—that strand 2 could potentially pair with strand 3 as well as with strand 1[33] (Fig. 4, right). This led to the suggestion that alternative secondary structures in the *trp* leader transcript might provide the termination or read-through signals that RNA polymerase recognizes.

We could imagine, as shown in Fig. 5, that the position of the ribosome on the transcript determines which one of the alternative RNA secondary structures would be formed[46,57]. For example, a ribosome stalled at the Trp codons or the Arg codon by a shortage of either amino acid would sterically block 1:2 pairing, thereby allowing strand 2 to pair with strand 3 as soon as strand 3 had been synthesized. By the time strand 4 had been synthesized its potential pairing partner, strand 3, would be base-paired and unavailable, preventing the formation of termination structure, 3:4. If 3:4 did not form and if, as we also believe (see next section), the 3:4 structure is the termination signal that RNA polymerase recognizes, then the transcribing polymerase would not stop at the attenuator. Consequently, the structural genes of the operon would be transcribed, leading ultimately to the biosynthetic correction of the tryptophan deficiency. If, however, a ribosome stalled at the Thr codon or reached the stop codon, strand 2 would be sterically prevented from pairing with strand 3 and thus 3:4 would be free to form and signal transcription termination. If the ribosome stalled at the Gly codon strand 1 would pair with strand 2, thereby preventing 2:3 pairing and facilitating pairing of 3 with 4. According to this scheme it is the location of the ribosome on the transcript of the *trp* leader region that determines which alternative RNA secondary structure forms. This, in turn, directs RNA polymerase to terminate transcription or to read through

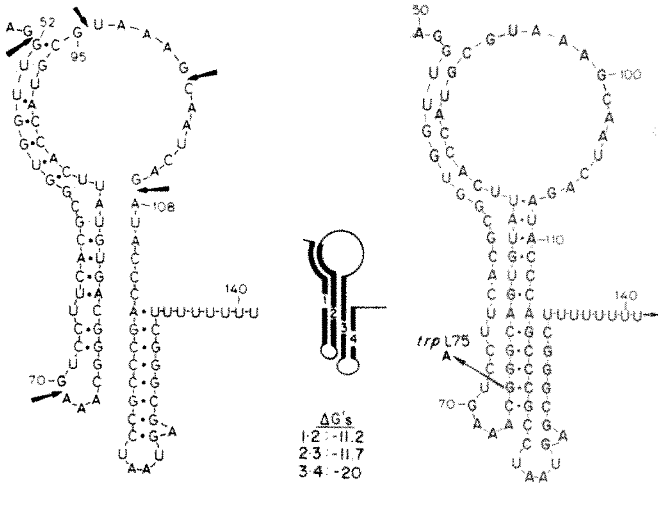


Fig. 4 Secondary structure alternatives in the *trp* leader transcript. In the large structure on the left the two base-paired structures that are detected *in vitro* are presented. The arrows indicate the sites of RNase Tl attack. The G-bonds in the hydrogen-bonded regions are not cleaved, presumably because the Gs are base paired. In the large structure on the right an alternative secondary structure is shown. Formation of this structure is thought to prevent transcription termination at the attenuator. For simplicity we designate the transcript segments that participate in these hydrogen-bonded structures strands 1–4 (see centre inset). The *trp* codons are in strand 1. The *trpL75* mutation is indicated.

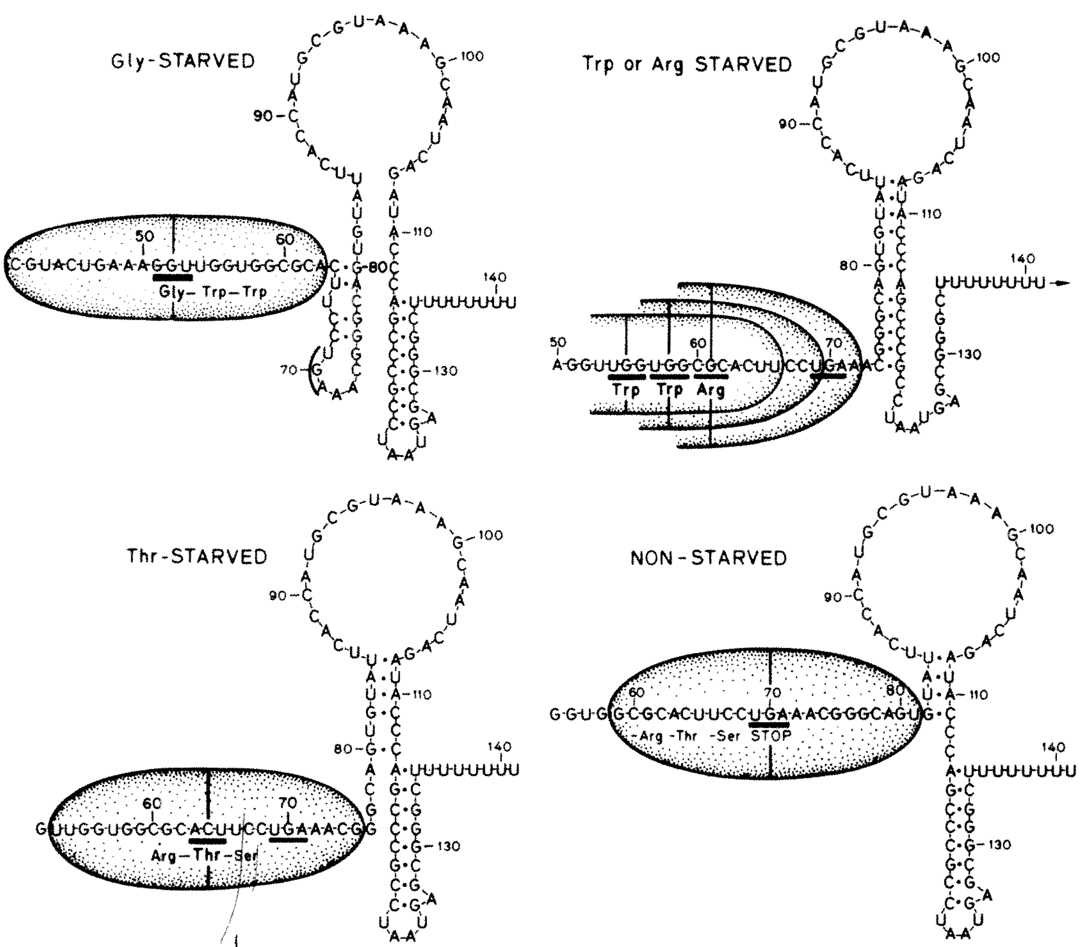


Fig. 5 Schematic representation of ribosome position on the *trp* transcript in cells starved of Gly, Trp, Arg or Thr, or cells that are not starved. It is proposed that a ribosome stalled over one of the codons of the transcript sterically blocks about 10 nucleotides on the downstream segment of the transcript, thereby determining which alternative secondary structure will form during transcription. Only when cells are starved of Trp or Arg is the ribosome positioned so that structure 2:3 will form. The existence of structure 2:3 could prevent pairing of 3 with 4 while 4 is being synthesized. It is assumed that transcription terminates at the attenuator only when structure 3:4 is formed during transcription.

into the structural genes of the operon. With this model for reference we can explain why the Arg-Thr-Ser codon sequence that follows the tandem Trp codons of *trp* transcripts is conserved in the enterobacteria. The Arg-Thr-Ser coding region is within the segment of strand 1 that can base pair with strand 2 (Figs 4, 5). Conservation of this region presumably reflects a requirement for this pairing capability.

Mutations in the leader region of the *trp* operon

Two classes of termination-defective *trp* leader mutants have been isolated. One type terminates transcription at less than the normal frequency and therefore operon expression is increased[58,59]. The base-pair changes in approximately 30 mutants of this type have been determined and, with one exception, all are in the segment of the leader region that corresponds to the 3:4 base-paired RNA structure (Fig. 6). Each of these base-pair changes reduces the predicted stability of the 3:4 structure by about 50%. All of these mutations affect the central G·C base pairs of the structure (Fig. 6), which of course have the greatest effect on stability of the 3:4 structure. *In vivo*, each of these mutations results in a two- to fourfold increase in operon expression. When restriction fragments bearing these mutations are transcribed, 40–70% of RNA polymerase molecules read through the attenuator. With a wild-type template only 5% do so. Clearly, these single mutant changes markedly influence the termination of polymerase activity. Largely on the basis of these findings we believe that the hydrogen-bonded 3:4 structure is the termination signal that the transcribing polymerase recognizes. Consistent with this conclusion is the additional observation[33] that when transcription is carried out *in vitro* with ITP substituting for GTP, transcription is no longer terminated at the attenuator, presumably because I = C base-paired RNA is less stable than G ≡ C base-paired RNA. Similar *in vitro* transcription studies with other base analogues that affect RNA secondary structure also suggest that a stable 3:4 RNA secondary structure is necessary for transcription termination[60].

A role for the A + T-rich as well as the G + C region of the attenuator in termination of transcription is suggested by our discovery of a termination-defective *trp* leader mutant with a single base-pair change in the A + T-rich region of the attenuator[58]. In addition, a deletion mutant exists in which the distal four A + Ts of the A + T-rich region are replaced by a foreign sequence[31]. This change eliminates transcription termination at the attenuator *in vitro* and reduces it *in vivo*[31,60]. Certain mutants altered in the β subunit of RNA polymerase are also aberrant in termination of transcription at the *trp* attenuator (C. Yanofsky and V. Horn, unpublished). In one class termination is partially relieved, while in a second, termination is more frequent. These mutations influence the termination of transcription only in strains which can form the 3:4 base-paired structure, strengthening the conclusion that this structure is the termination signal that is recognized by RNA polymerase.

There are also mutations in the *trp* leader region which increase termination at the attenuator *in vivo*[53]. Some mutations of this type have a second interesting and important property; they prevent the relief from termination that is normally associated with tryptophan or arginine starvation[53]. This observation provides the strongest evidence for our model of attenuation. In one of these mutants, *trpL29*, which has been isolated on two separate occasions, the AUG start codon for the leader peptide is replaced by AUA, a change which would prevent efficient translation of the leader transcript. Since strains with this mutation do not respond to tryptophan starvation by relieving termination, a prediction of our model—that a ribosome must move to the Trp codon region of the transcript for transcription termination to be relieved—is supported. A second type of mutant, *trpL75* (Fig. 4), of which there have also been two isolates, has a G → A change at position 75 in the leader region. Since this change essentially eliminates the pairing of strand 2 with strand 3, thereby leading to the formation of the 3:4

termination structure, our model explains why termination occurs despite tryptophan starvation. These considerations support two important features of our model: first, that a ribosome must translate the transcript at least to the first Trp codon for termination relief, and second, that pairing of strand 2 with strand 3 is also required for this relief. Thus, these observations strongly suggest that alternative secondary structures in the RNA transcript determine whether or not termination will occur at the *trp* operon attenuator.

Why do these two types of mutants exhibit increased termination of transcription at the attenuator *in vivo*? In the wild-type bacterium we assume that operon expression provides for some transcription readthrough, even in the presence of excess tryptophan[58]. This could be due to rapid dissociation of the translating ribosome from the stop codon of the transcript, allowing the 2:3 alternative structure to form. If this happens, formation of the 3:4 structure would be prevented and the transcribing RNA polymerase molecule would continue into the structural genes of the operon. This model can explain the increased termination observed *in vivo* in the two types of termination mutants. The AUA translation initiation-defective mutant would, during transcription, form structure 1:2. This would prevent pairing of strand 2 with 3, thereby promoting the formation of termination structure 3:4. In mutant *trpL75*, pairing of strand 2 with strand 3 is presumably defective. Hence, despite ribosome dissociation from the stop codon, strand 2 could not pair with strand 3, and therefore 3:4 would form more often. Thus, the increased termination of transcription observed in these mutants can be readily accommodated within our model.

Also consistent with our model is the observation that neither mutant change affects transcription termination at the attenuator *in vitro*[58].

The leader pause site

When restriction fragments containing the *trp* promoter are transcribed we observe the 140-nucleotide terminated leader transcript as the major product along with a small amount of a readthrough transcript. However, at early times during a single round of transcription, a third, major transcript is seen (M. Winkler and C. Yanofsky, unpublished). This species—which we call the pause transcript—is 90 nucleotides long and has at its 3′ terminus the sequence that could form the 1:2 base-paired structure (Fig. 4). Enhancement of pausing at this site when BrUTP replaces UTP (P. J. Farnham and T. Platt, unpublished) supports this proposal since numerous U's are present in the 1:2 stem. Kinetic studies indicate that every polymerase molecule that transcribes the initial segment of the leader region pauses at the 'leader pause site'. Eventually, these polymerase molecules resume transcription so that the pause transcript grows to become either the terminated transcript or the readthrough transcript. The position of the pause site in the leader region suggests that it could be used to synchronize ribosome attachment, movement and positioning on the leader transcript relative to the transcribing RNA polymerase molecule. The polymerase molecule might be held at the pause site until the translating ribosome reached a particular position on the transcript. Thereafter, the polymerase molecule would resume transcription and the translating ribosome would reach the positions on the transcript that favour one or the other of the alternative RNA secondary structures.

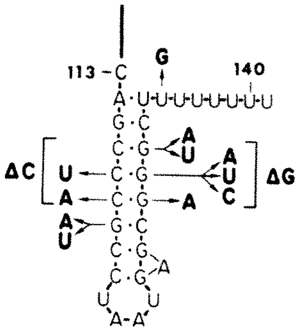


Fig. 6 Base pair changes in the *trp* leader region that relieve transcription termination at the attenuator. The changes in the transcript are shown; all are in the 3:4 structure and reduce its stability except the single U→G change at position 135.

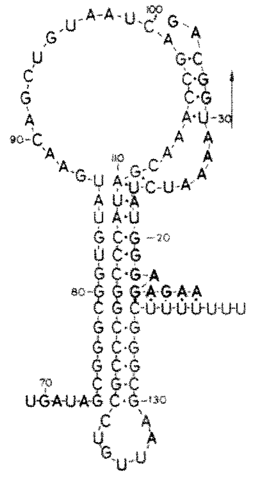


Fig. 7 Pairing of the leader ribosome binding site with the distal portion of the *trp* leader transcript of *S. typhimurium*. Note that both the Shine–Dalgarno region and the start codon region are hydrogen bonded to the distal segment of the transcript. Pairing of the indicated RNA segments has been demonstrated experimentally (K. Brown and C. Yanofsky, unpublished).

The leader ribosome binding site

Perusal of the sequences of the *trp* leader transcripts of various organisms and the results of secondary structure studies with these transcripts suggest an additional feature of the attenuation process. When the secondary structures formed in *trp* leader transcripts from different enterobacteria were analysed, another base-paired species was detected in addition to the 1:2 and 3:4 structures (K. Brown and C. Yanofsky, in preparation). This species contains an early segment of the transcript—the region containing the leader ribosome binding site—base-paired to a distal segment (Fig. 7). The calculated free energy for this base-paired segment is about $-32\ \mathrm{kcal\ mol^{-1}}$. If this pairing occurs *in vivo* the distal segment of the transcript could block the leader ribosome binding site, thereby preventing additional rounds of translation. Thus, once the decision was made to terminate transcription or to allow readthrough at the *trp* attenuator, additional translation of the peptide coding region would be prevented by masking the leader ribosome binding site. Sequence arrangements that allow pairing with the leader ribosome binding site may be a common feature of the structure of transcripts of operons regulated by attenuation (ref. 54 and M. Blumenberg and C. Yanofsky, unpublished). This situation is reminiscent of the masking of ribosome binding sites in RNA phages[61,62].

Summary of the mechanism of attenuation

Our current view of the events accompanying attenuation in the *trp* operon is as follows: a RNA polymerase molecule that escapes repressor control initiates transcription of the operon.

Soon after the segment of the transcript containing the leader ribosome binding site has been synthesized a ribosome attaches. Synthesis of the leader peptide begins while transcription of the later regions continues. The transcribing RNA polymerase molecule reaches the transcription pause site where it remains briefly, allowing the translating ribosome to approach. When RNA polymerase resumes transcription, the translating ribosome either stalls at one of the Trp codons, if the cell is deficient in tryptophan, or, if tryptophan is plentiful, moves to the translation stop codon. The position of the ribosome on the transcript determines which alternative RNA secondary structure forms. Only if structure 3:4 forms does the polymerase terminate transcription at the attenuator. Whether termination occurs or not, the leader ribosome binding site is then blocked by pairing with a transcript segment.

The interpretation given above presumes that changes in the secondary structure of transcripts mediate the regulatory response. This has not been proved! Structural studies on transfer RNAs have established that interactions other than hydrogen bonding contribute to three-dimensional form. Furthermore, a regulatory role for the DNA template has not been excluded. We should therefore keep open the possibility that other interactions also participate in establishing the informational transcript or template structures that are crucial in attenuation.

Attenuation in other amino acid biosynthetic operons

Extensive regulatory studies have been performed with the *his* operons of *S. typhimurium* and *E. coli*. This operon was the first gene cluster for which there was convincing evidence suggesting that charging of a transfer RNA had a regulatory role[8]. The leader regions of the *his* operons of *S. typhimurium* and *E. coli* have been sequenced[63,64] and appear to have structures resembling that of the *trp* operon. A potential ribosome binding site is situated in the leader transcript sequence, which, if used, would give a 16-residue peptide containing 7 adjacent His residues. A G+C-rich region with dyad symmetry is located somewhat beyond and is followed by an A+T-rich segment within which transcription termination probably occurs. The transcript segment corresponding to the G+C-rich region could form a stable stem and loop that could serve as a termination signal. A potential alternative competing secondary structure can be drawn for the transcript that could prevent formation of the presumed termination structure, depending on the location of the ribosome translating the leader peptide-coding region[63–65]. *In vitro* studies have implicated translation in the regulation of transcription of the *his* operon[66]. In addition, analyses with regulatory mutants altered at key positions in the *his* leader region provide insight into the parts played by different segments of this region or more likely, the corresponding transcript[65]. An ochre nonsense mutation in the leader peptide-coding region decreases operon expression to the extent that the corresponding mutant is a His auxotroph. This observation suggests that translation of at least the initial part of the leader peptide coding region is essential for operon expression. Suppression of the ochre mutation increases *his* enzyme levels and restores the His$^+$ phenotype. A second mutation conferring His auxotrophy occurs just beyond the translation stop codon for the leader peptide. Oddly enough, although this mutation does not introduce an amber codon, it is suppressed by amber suppressors. Apparently, the mutation prevents formation of the competing RNA structure, resulting in more efficient termination at the attenuator[65]. Presumably, the amber suppressors increase operon expression by allowing the translating ribosome to translate the leader peptide stop codon and disrupt the transcript termination structure[65]. One verified prediction of this interpretation was that amber suppressors would increase expression of the wild-type *his* operon.

Finally, a deletion that removes part of the termination structure increases operon expression. Thus, all aspects of attenuation in the *his* operon are similar to those of the *trp* operon, suggesting that attenuation is used in much the same way.

The *thrA$_1$A$_2$BC* and *ilvGEDA* operons of *E. coli* have an additional feature readily accommodated by the attenuation mechanism regulating amino acid biosynthetic operons[67–69]. Expression of these two operons is influenced by two and three amino acids, respectively (multivalent control). Obviously, if a leader transcript contains multiple codons for two or three amino acids in the critical pairing regions, the operon could respond to starvation for any one of these amino acids. The sequences of the leader regions of the *thr* and *ilv* operons suggest that this is exactly what occurs. The *thr* leader region encodes a peptide rich in Thr and Ile residues (Fig. 3); it contains a G+C-rich, A+T-rich region resembling other terminators, and its sequence predicts the formation of alternative, competing paired transcript structures that would be influenced by ribosome position on the transcript[67]. Mutants establish that the G+C-rich region is essential for the termination of transcription. The *ilvGEDA* operon is a more extreme example of multivalent control. It is regulated by changes in the extent of charging of Val, Ile and Leu tRNAs[8]. The leader region codes for a putative 32-residue peptide with multiple Ile, Val and Leu residues and contains adjacent G+C- and A+T-rich regions at which transcription is terminated *in vitro*[68,69]. Mutually exclusive secondary structures have been predicted for the transcript. Their role in the termination of transcription is supported by the finding that substitution of ITP for GTP in the *in vitro* transcription reaction prevents termination. The potential secondary structures in the transcript are complex, perhaps reflecting the need to respond to starvation for any one of the three amino acids. Stalling of more than one ribosome on the transcript may be required for maximum relief of termination[68].

The *leu* operon of *S. typhimurium* has a leader region containing all the features found in the other amino acid operons[70,71]. A 160-nucleotide terminated transcript is produced *in vitro*. The transcript codes for a 28-residue peptide containing 4 contiguous Leu residues. Alternative secondary structures can be predicted for the transcript, including a termination structure, a competing structure and a third structure that could prevent formation of the competing structure if translation of the transcript stopped early in the peptide-coding region. Again, substituting ITP for GTP in the *in vitro* transcription reaction prevents termination.

The Leu codons of the transcript are so positioned that ribosome stalling in the Leu codon region would promote formation of the RNA structure that is the alternative to the termination structure.

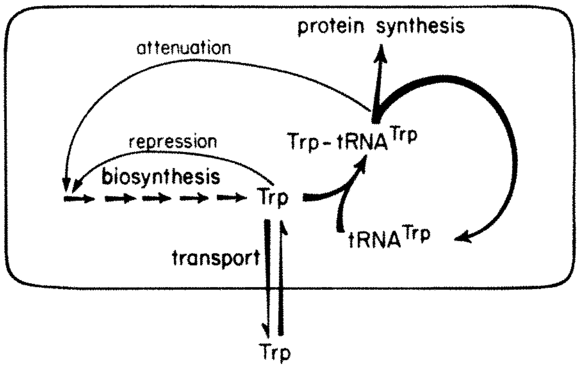


Fig. 8 Metabolic processes involved in tryptophan synthesis and utilization. The existence of the two transcription control mechanisms for the *trp* operon allows the bacterium to sense both the intracellular concentration of tryptophan and the extent of charging of tRNA^{Trp}.

The *pheA* operon of *E. coli* is a single gene operon containing a leader region with all the features seen in other amino acid operons controlled by attenuation[72]. A terminated transcript is produced *in vitro* and the site of termination is a typical termination region. The predicted *phe* leader peptide is 15 residues in length and contains 7 Phe residues in the last two-thirds of the peptide. Stable RNA secondary structures can be drawn for the *phe* transcript. Ribosome stalling over the Phe codons of the transcript could prevent formation of the transcript termination structure.

Thus, despite considerable differences in the sequences of leader regions of the *trp*, *his*, *thr*, *ilv*, *leu* and *phe* operons, the predicted structures of each of their transcripts allows the same mechanism of attenuation. When we compare the specific elements of repression systems with those of attenuation as they are used in amino acid biosynthetic operons, it is obvious that from the point of view of information content attenuation is a more economical regulatory mechanism. The unique regulatory information of attenuation is provided by the sequence of 50 or so base pairs of the leader region of each operon. This information is sufficient to be interpreted in terms of the appropriate regulatory response by the transcriptional and translational machinery of bacterial cells.

Conclusions

The existence of two mechanisms for regulating transcription of the *trp* operon leads one to wonder what advantage it is to the organism to have this apparent duplication of regulatory effort. The two mechanisms clearly allow similar responses to the same physiological situations. A possible explanation can be offered by considering the various metabolic reactions involved in the biosynthesis and utilization of tryptophan (Fig. 8). The repression system, by responding to tryptophan exclusively, is tailored to titrate the intracellular concentration of this amino acid. This concentration depends on several factors: first, tryptophan availability in the environment of the bacterium and its entry into the cell via transport systems; second, the rate of tryptophan biosynthesis; and third, the utilization of tryptophan for protein synthesis. By gauging the intracellular concentration of tryptophan, repression is particularly well suited to regulate expression of the operon in response to changes in the availability of tryptophan from the environment. We believe this to be its primary purpose. Attenuation is influenced by the extent of charging of tRNA^{Trp}. The relative concentrations of charged and uncharged tRNA^{Trp} are determined by several factors: the intracellular tryptophan concentration, the amount and activity of tryptophanyl-tRNA synthetase, the amount of tRNA^{Trp} and, perhaps most importantly, the overall rate of protein synthesis. It is this last factor that we believe provides the best explanation for attenuation. If a cell becomes deficient in one of the other amino acids the tRNA^{Trp} pool will become fully charged, resulting in maximal termination at the attenuator. Conversely, if a cell suddenly begins to synthesize proteins rapidly and tryptophan is in short supply, the tRNA^{Trp} pool will be relatively uncharged and there will be little or no termination at the attenuator. Thus, the attenuation mechanism nicely regulates expression of the operon relative to the overall rate of protein synthesis. The combined action of the two regulatory mechanisms enables the bacterium to recognize and respond to the principal external and internal events relevant to expression of the operon.

If it is advantageous to have two transcription control mechanisms for the *trp* operon of *E. coli*, why is attenuation the sole transcription control mechanism used to regulate expression of other amino acid biosynthetic operons such as the *his*, *leu* and *thr* operons[65,67,70]? The explanation for this apparent contradiction, I believe, is provided by our recent finding that two other operons are regulated by the *trp* repressor. These are the *aroH* operon, specifying one of three enzymes that catalyse the initial reaction in the common pathway of aromatic amino acid biosynthesis[73], and the *trp* repressor operon, encoding the *trp* aporepressor[40,74]. In both cases, tryptophan-activated repressor binds to the operator–promoter region of the operon and inhibits transcription initiation. However, the operators in these three operons are at different locations in their respective promoters[40]. This finding is consistent with the hypothesis that these operators evolved independently. Thus, we might speculate that in an organism ancestral to *E. coli* and other enterobacteria, the *trp* repressor might have been used to regulate *aroH* expression while the *trp* operon was regulated by attenuation alone. Subsequently, by very few mutations, an operator may have evolved in the *trp* operon, thereby permitting additional control over transcription of this operon. We suggest, therefore, that two regulatory mechanisms now exist for controlling transcription of the *trp* operon of *E. coli* because it was advantageous to the organism to have evolved a *trp* repressor binding site in this operon so that it could exploit the availability of the *trp* repressor.

Extensions

Attenuation as used in the regulation of amino acid biosynthetic operons is only one of several ways in which gene expression could be controlled through the termination of transcription. Alteration or modification of RNA polymerase, as is presumed to occur when it interacts with the N protein of bacteriophage λ[75–78], or interference with the action of proteins participating in transcription termination[78], are additional attenuation mechanisms. Transcription termination sites could also be used simply to reduce transcription of a gene relative to a leader segment or a preceding gene in the same operon. The *trpR* operon, specifying the *trp* aporepressor, has a termination site in its leader region which may function in this manner[40]. Similarly, the termination site located between *rplL* and *rpoB* in the *rplJL–rpoBC* operon may be used to regulate expression of the two RNA polymerase structural genes relative to the ribosomal protein genes[79]. Inadvertent attenuation also occurs when translation is prematurely terminated within one of the structural genes of a polycistronic operon. Here, in what is termed translational polarity[80], the termination of transcription reduces the transcription of distal genes[81]. In addition to these examples, one could imagine many interactions that would alter the secondary structure of RNA and thereby influence termination. Transcriptional pausing[82–84] should also be considered a mechanism of attenuation. Pause sites could conceivably be used to set the rate of transcription of a gene or operon below that attainable by polymerase–promoter interaction alone. Pause sites may thus fix the maximum rates of transcription. Studies on pausing in phage T7 and in the *trp* operon suggest that pausing can serve different regulatory roles[83,84] (M. Winkler and C. Yanofsky, unpublished). Finally, attenuation may not be limited to bacteria and their viruses but may also occur in animal viruses. In vesicular stomatitis virus (VSV), transcription decreases in an oriented manner at or near the junctions of neighbouring genes, resulting in a cumulative reduction in distal gene expression[85]. In adenovirus and VSV transcription short leader transcripts are accumulated[86,87], an observation suggesting attenuation. It is apparent from the examples cited that transcriptional attenuation is accomplished in diverse ways. We presume that these mechanisms bring into regulatory focus those cellular processes and components that cannot readily participate in the mechanisms that control the initiation of transcription.

Perusal of the reference list will convince the reader that the model developed in this article was based on the work of a large number of collaborators. To each, I express my sincerest appreciation for sharing in this adventure. I also thank my current co-workers and Irving Crawford, Terry Platt, Howard Zalkin and George Stauffer for their helpful comments, and the NSF, the USPHS and the American Heart Association for their continuing support. C.Y. is a Career Investigator of the American Heart Association.

1. Taylor, K., Hradecna, Z. & Szybalski, W. *Proc. natn. Acad. Sci. U.S.A.* **57,** 1618–1625 (1967).
2. Luzzati, D. *J. molec. Biol.* **49,** 515–519 (1970).
3. Kourilski, P., Bourguignon, M. F. & Gros, F. in *The Bacteriophage Lambda* (ed. Hershey, A. D.) 647–666 (Cold Spring Harbor, New York, 1971).
4. Heinemann, S. F. & Spiegelman, W. G. *Cold Spring Harb. Symp. quant. Biol.* **35,** 315–318 (1971).
5. Roberts, J. W. *Nature* **224,** 1168–1174 (1969).
6. Jackson, E. & Yanofsky, C. *J. molec. Biol.* **76,** 89–101 (1973).
7. Kasai, T. *Nature* **279,** 523–527 (1974).
8. Brenchley, J. E. & Williams, L. S. *A. Rev. Microbiol.* **29,** 251–274 (1975).
9. Cortese, R. in *Biological Regulation & Development* (ed. Goldberger, R.) 401–432 (Plenum, New York, 1979).
10. Singer, C. E., Smith, G. R., Cortese, R. & Ames, B. N. *Nature new Biol.* **238,** 72–74 (1972).
11. Lewis, J. A. & Ames, B. N. *J. molec. Biol.* **66,** 131–142 (1972).
12. Martin, R. G., Ames, B. N. & Hartman, P. E. in *7th int. Congr. Biochem.* Abstr., 261–262 (Tokyo, 1967).
13. Cohen, G. & Jacob, F. *C.r. hebd. Séanc. Acad. Sci., Paris* **248,** 2490–92 (1959).
14. Shimizu, Y., Shimizu, N. & Hayashi, M. *Proc. natn. Acad. Sci. U.S.A.* **70,** 1990–1994 (1973).
15. Zubay, G., Morse, D. E., Schrenk, W. I. & Miller, J. H. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1100–1103 (1972).
16. Squires, C. L., Lee, F. & Yanofsky, C. *J. molec. Biol.* **92,** 93–111 (1975).
17. Rose, J. K., Squires, C. L., Yanofsky, C., Yang, H. L. & Zubay, G. *Nature new Biol.* **245,** 133–37 (1973).
18. Squires, C. L., Rose, J. K., Yanofsky, C., Yang, H. L. & Zubay, G. *Nature new Biol.* **245,** 131–33 (1973).
19. McGeoch, D., McGeoch, J. & Morse, D. *Nature new Biol.* **245,** 137–40 (1973).
20. Hiraga, S., Ito, K., Hamada, K. & Yura, T. *Biochem. biophys. Res. Commun.* **26,** 522–527 (1967).
21. Kano, Y., Matsushiro, A. & Shimura, Y. *Molec. gen. Genet.* **102,** 15–26 (1968).
22. Ito, K. *Molec. gen. Genet.* **115,** 349–353.
23. Camakaris, J. & Pittard, J. *J. Bact.* **107,** 406–414 (1971).
24. Imamoto, F. *Nature* **220,** 31–34 (1968).
25. Hiraga, S. & Yanofsky, C. *J. molec. Biol.* **79,** 339–49 (1973).
26. Baker, R. & Yanofsky, C. *J. molec. Biol.* **69,** 89–102 (1972).
27. Morse, D. E. & Morse, A. N. C. *J. molec. Biol.* **103,** 209–26 (1976).
28. Bertrand, K., Squires, C. & Yanofsky, C. *J. molec. Biol.* **103,** 319–37 (1976).
29. Bronson, M. J., Squires, C. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **70,** 2335–39 (1973).
30. Squires, C., Lee, F., Bertrand, K., Squires, C. L., Bronson, M. J. & Yanofsky, C. *J. molec. Biol.* **103,** 351–81 (1976).
31. Bertrand, K., Korn, L. J., Lee, F. & Yanofsky, C. *J. molec. Biol.* **117,** 227–47 (1977).
32. Lee, F., Squires, C. L., Squires, C. & Yanofsky, C. *J. molec. Biol.* **103,** 383–393 (1976).
33. Lee, F. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4365–4369 (1977).
34. Lee, F., Bertrand, K., Bennett, G. & Yanofsky, C. *J. molec. Biol.* **121,** 193–217 (1978).
35. Bertrand, K. & Yanofsky, C. *J. molec. Biol.* **103,** 339–49 (1976).
36. Yanofsky, C. *J. Am. med. Assoc.* **218,** 1026–1035 (1971).
37. Platt, T. in *The Operon* (eds Miller, J. & Reznikoff, W.) 268–302 (Cold Spring Harbor, New York, 1978).
38. Rose, J. K. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **71,** 3134–38 (1974).
39. Bennett, G. N. & Yanofsky, C. *J. molec. Biol.* **121,** 179–192 (1978).
40. Gunsalus, R. G. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
41. Oppenheim, D., Bennett, G. & Yanofsky, C. *J. molec. Biol.* **144,** 133–142 (1980).
42. Jackson, E. & Yanofsky, C. *J. molec. Biol.* **69,** 307–313 (1972).
43. Crawford, I. P. & Stauffer, G. V. *A. Rev. Biochem.* **49,** 163–195 (1980).
44. Yanofsky, C. & Soll, L. *J. molec. Biol.* **113,** 663–677 (1977).
45. Vold, B. S., Lazar, J. M. & Gray, A. M. *J. biol. Chem.* **254,** 7362–7367 (1979).
46. Eisenberg, S. P., Yarus, M. & Soll, L. in *Transfer RNA* (eds Soll, D., Schimmel, P. & Abelson, J.) (Cold Spring Harbor, New York, in the press).
47. Korn, L. & Yanofsky, C. *J. molec. Biol.* **103,** 395–409 (1976).
48. Korn, L. & Yanofsky, C. *J. molec. Biol.* **106,** 231–241 (1976).
49. Rosenberg, M. & Court, M. *A. Rev. Genet.* **13,** 319–353 (1979).
50. Platt, T., Squires, C. & Yanofsky, C. *J. molec. Biol.* **103,** 411–420 (1976).
51. Schmeissner, U., Ganem, D. & Miller, J. H. *J. molec. Biol.* **109,** 303–326 (1977).
52. Miozzari, G. F. & Yanofsky, C. *J. Bact.* **133,** 1457–1466 (1978).
53. Zurawski, G., Elseviers, D., Stauffer, G. V. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5988–5992 (1978).
54. Miozzari, G. F. & Yanofsky, C. *Nature* **276,** 684–689 (1978).
55. Miozzari, G. F. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5580–5584 (1978).
56. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. *Nature* **272,** 414–422 (1978).
57. Oxender, D., Zurawski, G. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5524–5528 (1979).
58. Zurawski, G. & Yanofsky, C. *J. molec. Biol.* **142,** 123–129 (1980).
59. Stauffer, G. V., Zurawski, G. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4833–4837 (1978).
60. Farnham, P. J. & Platt, T. *Cell* **20,** 739–748 (1980).
61. Lodish, H. F., in *RNA Phages* (ed. Zinder, N. D.) 301–318 (Cold Spring Harbor, New York, 1975).
62. Fiers, W. in *RNA Phages* (ed. Zinder, N. D.) 353–396 (Cold Spring Harbor, New York, 1975).
63. Di Nocera, P. P., Blasi, F., DiLauro, R., Frunzio, R. & Bruni, C. B. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4276–4280 (1978).
64. Barnes, W. M., *Proc. natn. Acad. Sci. U.S.A.* **75,** 4281–4285 (1978).
65. Johnston, H. M., Barnes, W. M., Chumley, F. G., Bossi, L. & Roth, J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 508–512 (1980).
66. Artz, S. W. & Broach, J. R. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3453–3457 (1975).
67. Gardner, J. F. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1706–1710 (1979).
68. Lawther, R. P. & Hatfield, G. W. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1862–1866 (1980).
69. Nargang, F. E., Subrahmanyam, C. S. & Umbarger, H. E. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1823–1827 (1980).
70. Gemmill, R. M., Wessler, S. R., Keller, E. B. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4941–4945 (1979).
71. Keller, E. B. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6186–6190 (1979).
72. Zurawski, G., Brown, K., Killingly, D. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4271–4275 (1978).
73. Zurawski, G., Gunsalus, R. P., Brown, K. D. & Yanofsky, C. *J. molec. Biol.* (in the press).
74. Singleton, C. K., Roeder, W. D., Bogosian, G., Somerville, R. L. & Weith, H. L. *Nucleic Acids Res.* **8,** 1551–1560 (1980).
75. Adhya, S., Gottesman, M. & deCrombrugghe, B. *Proc. natn. Acad. Sci. U.S.A.* **71,** 2534–2538 (1974).
76. Franklin, N. C. *J. molec. Biol.* **89,** 33–48 (1974).
77. Gottesman, M. E., Adhya, S. & Das, A. *J. molec. Biol.* **140,** 57–75 (1980).
78. Greenblatt, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 1991–1994 (1980).
79. Barry, G., Squires, C. L. & Squires, C. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4922–4926 (1979).
80. Adhya, S. & Gottesman, M. *A. Rev. Biochem.* **47,** 967–996 (1978).
81. Richardson, J. P., Grimley, D. & Lowery, C. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1725–1728 (1975).
82. Gilbert, W. in *RNA Polymerase* (eds Losick, R. & Chamberlin M.) 193–205 (Cold Spring Harbor, New York, 1976).
83. Kassavetis, G. A. & Chamberlin, M. J. *J. biol. Chem.* (in the press).
84. Kingston, R. E., Nierman, W. C. & Chamberlin, M. J. *J. biol. Chem.* (in the press).
85. Iverson, L. E. & Rose, J. K. *Cell* (in the press).
86. Fraser, N. W., Sehgal, P. B. & Darnell, J. E. Jr. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2571–2575 (1979).
87. Testa, D., Chanda, P. K. & Banerjee, A. K. *Cell* **21,** 267–275 (1980).

ARTICLES

VLBI structures of the images of the double QSO 0957+561

R. W. Porcas

Max-Planck-Institüt für Radioastronomie, Auf dem Hügel 69, Bonn, FRG

R. S. Booth, I. W. A. Browne, D. Walsh & P. N. Wilkinson

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

Very long baseline interferometry (VLBI) observations of the double QSO 0957+561 have revealed radio fine structure in the two 'image' components A and B. The structures are similar, both of the 'core–jet' type typical of many compact extragalactic radio sources. The shapes of the images put strong constraints on the mass distribution responsible for the gravitational imaging.

THE twin QSOs 0957+561A, B are widely believed to be images of a single QSO produced by the gravitational lens effect[1]. The lens has been identified with an 18.5 mag elliptical galaxy in a distant cluster[2–5]. The similarity of the optical spectra of the images has been confirmed[6,7] and the flux ratio of the images has been shown to be the same over a wide range of frequencies covering the radio to UV parts of the spectrum[8–17].

The asymmetry of the geometry of the 'images' A and B with respect to the galaxy indicates that the lens action cannot be modelled by a spherical mass distribution[3]. The gravitational

0028-0836/81/090758—05$01.00

field of the surrounding galaxy cluster and the (non-spherical) mass distribution of the elliptical galaxy also determine the geometry of the images. In previous VLBI observations[10] we showed that the radio components A and B both have compact structure on a scale size of <20 marc s. We now report the results of more extensive and more sensitive VLBI observations at two frequencies, using inteferometers with resolutions ranging from 60 to 1 marc s.

Observations

We have made two sets of VLBI observations of the double QSO with the 2-MHz bandwidth Mark II recording system, using a transatlantic baseline and the large antennas available in the European VLBI network. On 9–10 January 1980 we observed at 4,996 MHz using the Max-Planck-Institüt für Radioastronomie (MPIfR) 100-m telescope at Effelsberg and the NRAO 43-m telescope at Green Bank, West Virginia. On 22–23 February we observed at 1,666 MHz using the Effelsberg and Green Bank antennae, and also the Jodrell Bank 76-m and Onsala 25-m antennae. Hydrogen maser frequency standards were used at Effelsberg, Onsala and Green Bank, and a rubidium standard at Jodrell Bank. All the cross-correlation of the signals for these experiments was done with the MPIfR processor in Bonn. In the subsequent analysis we attempted to achieve the highest possible sensitivity by integrating the cross-correlation function for as long as the coherence of the local oscillators, atmosphere and ionosphere would allow. This coherence time was determined by observation of the calibration sources 1642+69, BL Lac and 0235+16, which also served to calibrate the gain of the interferometers.

With the resolution used in these observations the double QSO consists only of two compact 'image' components A and B and there is no evidence of any compact structure associated with the extended components seen in lower-resolution maps. The A and B components are characterized by separate residual fringe rates in the cross-correlation function, representing their offsets from a reference 'fringe stopping' position. The long integration times then permit a clear separation of the visibility functions of A and B in the fringe rate spectrum for all but a few periods of time on each baseline (see ref. 10). Data corresponding to these periods were not used in the analysis.

Image structures

The highest sensitivity was obtained at 1,666 MHz. With 10-min integration times, the r.m.s. noise errors on the baselines Effelsberg–Jodrell, Onsala–Effelsberg and Onsala–Jodrell were 3, 6 and 6 mJy, respectively, with minimum lobe spacings in the range 30–60 marc s. On the Effelsberg–Green Bank baseline, which had resolution 6 marc s, the noise level was 5 mJy after 20 min integration. The visibility functions obtained for A and B on these baselines are shown in Fig. 1*a–h*.

The 1,666-MHz data show that A and B exhibit remarkably similar behaviour, both visibility functions indicating structure on the scale of ~50 marc s. The data for the baseline Effelsberg–Jodrell can be compared directly with that reported in Porcas *et al.*[10] which was taken at times corresponding to UT 16–19 h and 0.5 h in our present observations. These are all times which, by chance, were well away from the minima in the visibility functions and we were unable then to attribute the small amplitude differences to source structure.

To derive the source structures from the visibility functions we have fitted simple two-component gaussian models to the 1,666-MHz data. The data from both A and B are fitted well by models consisting of a barely resolved (core) component and an elongated (jet) component. The parameters of the best fitting models are given in Table 1. The visibilities predicted by the model are shown by the solid lines drawn through the data points in Fig. 1. Table 1 also shows errors where the parameter values are well defined by the data. This applies in particular to the position angle (PA) and separation of the jets from the cores in both A and B. Other parameters are less well defined because of strong coupling between their values, for example, the relative flux

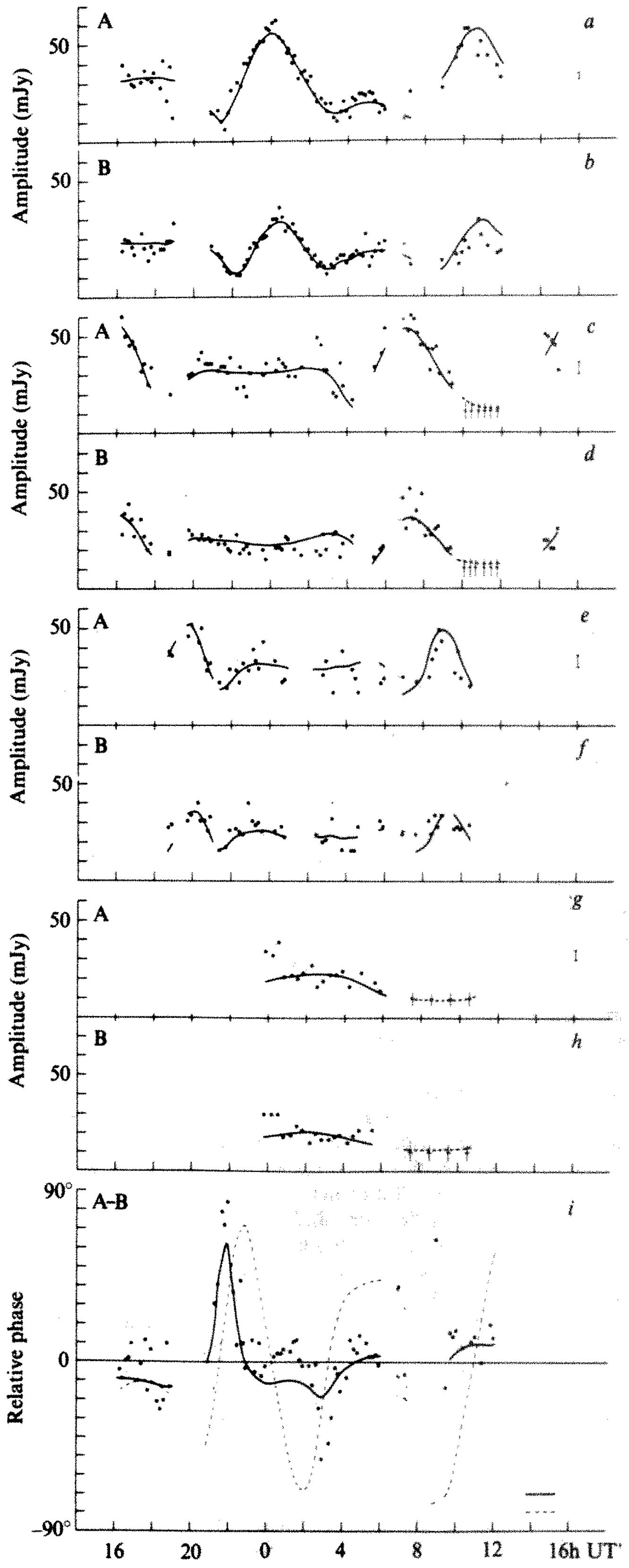


Fig. 1 *a–h*, Amplitudes of the visibility functions of components A and B at 1,666 MHz on the following baselines: *a,b*, Effelsberg–Jodrell Bank; *c,d*, Onsala–Effelsberg; *e,f*, Onsala–Jodrell Bank; *g,h*, Effelsberg–Green Bank. The solid line through the points represents the two-component gaussian models whose parameters are given in Table 1. 1 σ error bars are shown in the top right corners. *i*, Relative phase of the visibility functions (component A–component B) at 1,666 MHz on the baseline Effelsberg–Jodrell Bank. The solid line represents the relative phase predicted from the models fitted to the amplitude data for similarly oriented structures; the broken line is that predicted for opposite orientations. (At 0 h UT, Greenwich ST = 10h 08min.)

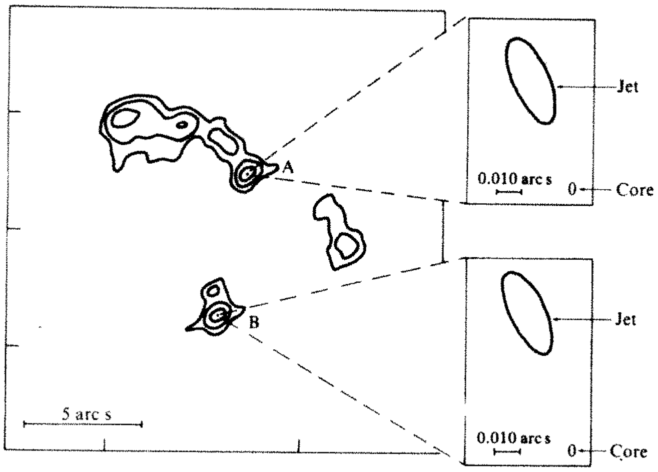


Fig. 2 Outline of the VLBI structures of the images A and B (half-power contour of the fitted gaussian models) projected against the VLA map of Greenfield *et al.*[12].

density between the cores and jets, and their sizes. We have further constrained the flux densities of the jets, however, by requiring that the total model flux density for each component does not exceed the corresponding total flux density derived from low-resolution maps. For this purpose we used the spectra given by Noble and Walsh[13] derived from published flux density values at various frequencies. This procedure seems appropriate as the maximum correlated flux densities we obtained equal these derived values (61 mJy and 43 mJy, respectively).

In general we have only used data from integration periods where we have unambiguously detected both A and B. Occasionally, however, both are undetected, in particular during a 4-h period at the end of the track on the Effelsberg–Green Bank baseline. We have used this information to constrain the models to give low values in these periods and its main effect is to require a slight extension of the core components of ~3 marc s in PA ~−10°. One region where our gaussian model parametrization of the structures is not quite adequate is at the beginning of the Effelsberg–Green Bank track. This indicates the existence of an additional fine-scale structure not represented by these simple core–jet models.

Model fitting to only the amplitudes of the measured visibility functions leads to a 180° ambiguity in the orientations of the structures in both A and B. To resolve this ambiguity we have analysed the relative visibility phase between components A and B on the baseline Effelsberg–Jodrell, which has the highest signal-to-noise ratio (and hence smallest phase noise errors). As the components A and B are separated by only ~6 arc s, measurements of their relative phase are effectively uncorrupted by the various perturbations caused by the atmosphere and the equipment. The relative phase, therefore, represents the true difference between the phase of the visibility functions of A and B. This itself is composed of (1) a sinusoidal function of period 24 h representing the position difference between A and B, and (2) other terms representing any difference in their structure.

We first removed from the measured phase difference a term corresponding to the centroid separation measured by Porcas *et al.*[10]. The correct values are $\Delta RA = -0.1499$ s, Δ dec $= +6.047$ arc s. Due to an algebraic error, the differential precession correction was wrongly applied for the right ascension difference quoted in our previous paper. The corrected values agree well with those quoted by Gorenstein *et al.*[19] from high-sensitivity transatlantic VLBI measurements at 2.3 GHz. The residual phase difference, representing the difference in the structures, is shown in Fig. 1*i*.

In interpreting Fig. 1*i* we consider the following points. If the structures of A and B were identical and oriented in the same sense, the difference phase would be everywhere zero. If A and B had oppositely directed structures then, as the individual structures are themselves asymmetric, the relative phase would reflect this asymmetry, giving values twice that of the phase of either individual visibility function. We assume that the measured centroid separation in fact corresponds to that of the core components in A and B. We would then expect to see these phase departures occurring at hour angles where the jet component (which is 'resolved out' at most hour angles where the projected baseline is not perpendicular to the jet) starts contributing to the correlated flux density.

For the case we have here, the structures of A and B differ anyway, primarily in the separations between the core and jet. Thus, even in the case of the same orientation, we expect imperfect cancellation of the visibility phases where the visibility functions differ most, that is near the minima. This is indeed apparent from Fig. 1*i*, where there are rapid relative phase excursions at UT 22h and 03h. Of particular interest, however, is the return of the relative phase to near zero immediately after the 22h excursion and lasting until the 03h excursion. This is precisely what would be expected for similar orientations as, when both jets contribute to the correlated flux densities, the centroid separation will be similar to that of the cores. (Indeed, this property has enabled us to deduce the correct centroid separation from the consistent phase behaviour away from the minima in our 1979 data.) This is not true for oppositely oriented structures.

This agreement is confirmed by examination of the relative phase predicted from the models of the A and B structures fitted to the amplitude data. The solid curve in Fig. 1*i* is that for similar orientation of structures, with both A and B jets pointing to the north-east; the dashed curve is that for opposite orientations. Although the models do not accurately reproduce the depth of the phase excursion, the gross features of the measured phase difference agree well for the case of the same orientation.

Table 1 Gaussian model parameters

	Separation (marc s)	PA	Flux density (mJy)		Maj. axis (marc s)	Axial ratio	PA
Image A							
Core	–	–	29	61 ± 3	3.5 ± 1.5	0.5 ± 0.1	−10° ± 6°
Jet	46 ± 1	21° + 1°	32		38 ± 3	0.35 ± 0.04	27° ± 3°
Image B							
Core	–	–	24	41 ± 2	2.5 ± 2.7	0.5 ± 0.2	−7° ± 15°
Jet	56 ± 2	17° ± 1°	17		37 ± 6	0.36 ± 0.1	28° ± 3°

A more indirect way of obtaining the component orientations individually uses the 'closure' phase, that is the algebraic sum of three simultaneously measured visibility phases around a closed triangle of baselines[18]. However, to obtain useful closure phase information the data must have a good signal-to-noise ratio on all three baselines for a sufficient period. This was not the case for our data for not only are those from the Onsala baselines very noisy but also there is only one period, from 21.30 to 00.40 UT, where we have reliably separated the contributions from A and B on all three baselines simultaneously. The closure phases for this 3-h period do show features corresponding to the rapid phase changes occurring on the Effelsberg–Jodrell baseline during this time; however, although these are consistent with our deductions from the relative phase, the closure phases do not in themselves provide a reliable independent check on the component orientations.

Our highest-resolution data were obtained from the Effelsberg–Green Bank baseline at 4,996 MHz, with minimum lobe spacing of 2 marc s. With a 20-min integration time we obtained an r.m.s. noise level of 8 mJy, and for much of the track we detected both A and B components at levels of ~28 and ~25 mJy respectively, indicating very compact (≤ 1 marc s) structure in the cores. For the later part of the track, however, both components get weaker and fall below our detection limit. This suggests that the compact core components of A and B are slightly elongated at 4,996 MHz in the same direction as at 1,666 MHz. Haschick *et al.*[14] have obtained very similar data at 5,010 MHz on this baseline, although they attribute the non-detection of A and B at the end of the track to telescope pointing errors.

The agreement of the flux densities detected at the start of the track with the total 5-GHz flux densities of A and B shown by the lower-resolution maps[8,9,11,12] suggests that the jets seen at 1,666 MHz are much weaker at 5 GHz and thus have a steep spectrum.

Interpretation and discussion

The similarity of the structures of A and B clearly confirms that we are seeing two images of a single structure. The core–jet configuration seen in each image is typical of the compact structures of many QSOs and radio galaxies[20]. Figure 2 shows an outline of the model VLBI structures for A and B projected against a map of the larger-scale structure. In the lower-resolution maps at 5 GHz (ref. 12) and 408 MHz (ref. 13) there is an extended bridge of emission between A and the outermost extended structure to the north-east. The jet in the VLBI structure does not point directly at this bridge but it is possible to draw a smooth curve from the core of A through the jet and along the bridge suggesting a continuity between these features. Young *et al.*[4] suggest that the curvature of the complex consisting of A together with the extended structure to the north-east and to the west may be the result of gravitational imaging of a linear structure. The curvature implied by the VLBI structure seems too extreme for this explanation, and the structure of the object itself is probably non-linear. In fact, if component B is excluded, the overall structure is rather similar to that of the QSO 3C309.1 (ref. 21) which is also a triple source showing a marked rotation in position angle between the nucleus and the outer lobes.

The details of the VLBI structure place severe constraints on the mass distribution of the lens. The results of Young *et al.*[3] show that the previously observed properties of the double QSO cannot be produced by a single spherically symmetric mass distribution; such a distribution would require A, B and the galaxy G1 to be co-linear as opposed to the observed skew image line. In their investigation of mass distributions to account for previous observations, Young *et al.*[4] "found it remarkably difficult to fit both the skew in the line and the high brightness of B relative to A. Fitting either one is reasonably easy, but fitting both strained the capacity of the present models".

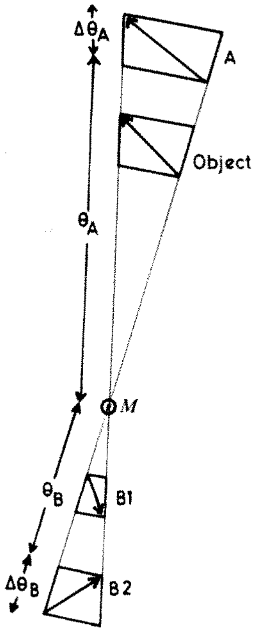


Fig. 3 Schematic representation of the imaging of an object element by a mass distribution, *M*, which has spherical symmetry. The arrows indicate an arbitrary line in the object and the images of this line.

Paradoxically the close similarity between the image structures derived from our VLBI observations is most surprising. To see why this should be so, it is instructive to compare them with those expected with a spherical mass distribution. Bourassa and Kantowski[22] showed that a spheroidal mass distribution will produce one or three images (or, more generally, an odd number). Thus we expect at least one additional image besides A and B. This has been discussed in the context of 0957 + 561 (refs 3, 23) and it is expected that the B image may be split into two separate images. We adopt the notation of Young *et al.*[3] and refer to these as B2 (the outermost image) and B1. Figure 3 illustrates the imaging of an element of the object. We can confidently identify our B image with B2 since the jets in both A and B components point north. Two properties of our model that are quite different from what is produced by a spherical mass distribution are (1) the relative PAs of the core–jet lines, and (2) the relative linear dimensions of the two images.

(1) With a spherical distribution the PAs of the core–jet lines should have opposite signs with respect to the line joining the two images and the galaxy. In our model this is not so; in fact the two lines are almost parallel. Referring to the actual skew galaxy–image configuration[3], the core–jet line of image A has PA +37° with respect to the line A–G1, and the core–jet line of image B has PA +6° with respect to the line B–G1. (We use Stockton's position[5] for G1, that is, 0.19″E and 1.00″N of B.) The fact that the line of structure B points almost directly at G1 must be attributed to chance, because it is determined by the corresponding line of the object which must have a random orientation if the gravitational imaging interpretation is correct.

(2) Because the surface brightness of the images and object must be the same, the flux density ratios of the two images produced by a spherical mass distribution is

$$\frac{S_A}{S_B}=\left|\frac{\theta_A}{\theta_B}\cdot\frac{\Delta\theta_A}{\Delta\theta_B}\right|$$

The ratio of the total flux densities is S_A/S_B is 1.42. (The ratio from our models is 1.9 for the jets and 1.2 for the cores. Though differential amplification is possible along the lengths of the images, these observed values are subject to the uncertain division of flux between core and jet in each component, as pointed out earlier; they cannot be regarded as significantly different.) The observed ratio $|\theta_A/\theta_B|=5$, so the value of $|\Delta\theta_B/\Delta\theta_A|$ from the above formula is 3.5. The best estimate of $|\Delta\theta_B/\Delta\theta_A|$ from our model comes from the core–jet separations and is 1.2 (this is confirmed by the major axes of the jets, for which a value 1.0 is obtained; taking into account the errors, this is not significantly different from 1.2). Thus the length ratio of the two images differs by a factor of 3 from that expected from the simple case of a spherical mass distribution.

It is of interest to put limits on the intensity of image B1. Our visibility functions for B are sensitive to structure on scales up to ~200 marc s. From the lack of any further modulation on our best defined visibility function (Effelsberg–Jodrell Bank) we can put an upper limit of ~5 mJy on the flux density of B1. This rules out Young *et al.*'s[3] model (c) for the system. For separations of B1 from B2 greater than this we become increasingly less sensitive due to smearing of the visibility function by our 10-min integration time, but the VLA map of Greenfield *et al.*[12] already puts strong limits on B1 in this angular range. There is, however, a weak compact feature on the VLA map, coincident with the galaxy G1 within the observational errors, and it has been suggested that this is either the image B1 or emission from G1 itself. This feature is too weak to be detected by our observations, so we cannot say whether it, too, is compact like the A and B components.

Finally, we consider the effect of any intrinsic variations in the radio and optical emission from the quasar. The estimation of the time delay in the variation between images A and B depends on detailed modelling of the lens galaxy and the surrounding cluster[3,4,24]. Variations of the relative intensity in the optical continuum emission have already been reported[4,25] and also in the 5-GHz radio emission[12] and the UV continuum (R. Wilson, personal communication). At frequencies of 1.6 GHz and lower much of the radio emission comes from the steep spectrum jet. Because this is extended and lies $\sim$100 pc away from the core component, we cannot expect this radio emission to vary on time scales as short as a few years. We would thus expect the ratio of the intensities of A and B to vary less at lower frequencies.

Our VLBI measurements of the structures of A and B therefore confirm the interpretation of 0957+561 as a QSO seen through a gravitational lens. The intrinsic radio structure of the QSO has two outer lobes and a compact central component consisting of a jet and a core. We have put limits on the properties of any third image which an extended, transparent lens is expected to produce. The surprisingly close similarity of the orientations of the images and their sizes must be attributed to the non-spherically symmetric elements of the imaging mass distribution.

We thank Professor I. Shapiro for pointing out our differential precession error in ref. 10, and Herr H. Blaschke and Herr Stursberg for operating the MPI VLBI processor. We also thank the staff at NRAO and Onsala for observations. NRAO is operated by Associated Universities Inc. under contract with the NSF. Onsala Space Observatory, Chalmers University of Technology, is operated with financial support from the Swedish NRC.

Received 31 October; accepted 29 December 1980.

1. Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381 (1979).
2. Gunn, J. E., Kristian, J., Oke, J. B., Westphal, J. A. & Young, P. J. *IAU Circ.* No. 3431 (1979).
3. Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. *Astrophys. J.* **241**, 507 (1980).
4. Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. *Astrophys. J.* (in the press).
5. Stockton, A. Preprint (University of Hawaii, 1980).
6. Weymann, R. J. *et al. Astrophys. J. Lett.* **233**, L43 (1979).
7. Wills, B. J. & Wills, D. *Astrophys. J.* **238**, 1 (1980).
8. Pooley, G. G. *et al. Nature* **280**, 461 (1979).
9. Roberts, D. H., Greenfield, P. E. & Burke, B. F. *Science* **205**, 894 (1979).
10. Porcas, R. W., Booth, R. S., Browne, I. W. A., Walsh, D. & Wilkinson, P. N. *Nature* **282**, 385 (1979).
11. Greenfield, P. E., Roberts, D. H. & Burke, B. F. *Science* **208**, 495 (1980).
12. Greenfield, P. E., Burke, B. F. & Roberts, D. H. *Nature* **286**, 865 (1980).
13. Noble, R. G. & Walsh, D. *Nature* **288**, 69 (1980).
14. Haschick, A. D., Moran, J. M., Reid, M. J., Davis, M. Lilley, A. E. Preprint (Center for Astrophysics, 1980).
15. Soifer, B. T. *et al. Nature* **285**, 91 (1980).
16. Lebofsky, M. J., Rieke, G. H., Walsh, D. & Weymann, R. J. *Nature* **285**, 385 (1980).
17. Gondhalekar, P. M. & Wilson, R. *Nature* **285**, 461 (1980).
18. Rogers, A. E. E. *et al. Astrophys. J.* **193**, 293 (1974).
19. Gorenstein, M. V. *et al. Bull. Am. astr. Soc.* **12**, 498 (1980).
20. Readhead, A. C. S., Cohen, M. H., Pearson, T. J. & Wilkinson, P. N. *Nature* **276**, 768 (1978).
21. Kus, A. J., Wilkinson, P. N. & Booth, R. S. *Mon. Not. R. astr. Soc.* (in the press).
22. Bourassa, R. R. & Kantowski, R. *Astrophys. J.* **195**, 13 (1975).
23. Dyer, C. C. & Roeder, R. C. *Astrophys. J. Lett.* **238** L67 (1980).
24. Dyer, C. C. & Roeder, R. C. *Astrophys. J.* (submitted).
25. Miller, J. S., Antonucci, R. R. J. & Keel, W. C. *Nature* **289**, 153 (1981).

Halogen and phosphorus storage in the Earth

J. V. Smith

Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

Speculations on the spatial distribution of elements involve geochemical, geophysical and cosmochemical factors. From analyses of basalts and minerals in peridotites, it is found that: apatite with $\sim$*3 wt% F, and up to 1 wt% Cl and 0.003 wt% Br, is the principal reservoir of halogens in the upper mantle and mica with* $\sim$*0.5 wt% F and 0.05 wt% Cl is a second reservoir; apatite with* $\sim$*18 wt% P accounts for most of the phosphorus in the mantle, but is insignificant with respect to* $\sim$*0.7 wt% P in the core; surface reservoirs contain at least two-thirds to three-quarters of the Cl and Br, but only one-fifth of the F and <1% of the P and finally the Cl and Br contents of the whole Earth are* $<66\times10^{18}$ *and* $<17\times10^{16}$ *kg, using a new estimate of* $\sim10\times10^{19}$ *kg F.*

ALL cosmochemical models for the Earth involve less efficient capture of volatile elements than refractory ones from the solar nebula, but accretion processes[1] are complex. The controversial Ganapathy–Anders[2,3] (GA) model relates elements that would condense nearly together during simple progressive cooling under equilibrium. Chemical analyses of surface reservoirs, coupled with compositions for interior zones inferred from geophysical and geochemical data, were used to obtain a range of estimates of the bulk composition of the Earth[4]. Particularly important are estimates of the extent of transfer of magmatophile elements from mantle to crust by partial melting, and back again by foundering[5–7]. In spite of uncertainties in assignment of elements to reservoirs a general correlation can be implied[2,4,8–12] between volatility in the nebular gas and concentration in the whole Earth. Probable limits on the weight fractions of 14 to 20×10^{-9} U and 14 to 20×10^{-5} K are important.

Table 1 compares estimates of F, P, Cl, K and Br in near-surface reservoirs and whole-Earth models. The last row shows masses assuming that the bulk Earth has the same composition as carbonaceous meteorite of type C1. These masses are multiplied by 1.5 to account for overall depletion of volatile elements in the whole Earth (p. 351 of ref. 4). New analyses[14] of halogens in C1 meteorites are lower than most earlier analyses[13], especially for F, and a revised F/Cl ratio of 0.09 fits electron microprobe analyses of chlorapatites in meteorites (see ref. 15) and other bulk analyses[16,17].

Phosphorus should condense early in the 'metal' fraction of a nebular condensation model[2]. Any uncertainty in the relative fractions of early condensate, silicate and metal components in the Earth is here considered trivial. Phosphorus is strongly bound in minerals of meteorites, and should not be lost to vapour during collisions of planetesimals. If P accretes with the same efficiency as Fe from a solar nebula with C1-type composition, its mass must lie between 10×10^{21} and 14×10^{21} kg, and probably between 11×10^{21} and 13×10^{21} kg.

Potassium and F fall into the volatiles 1,300–600 K group for which the depletion factor with respect to 1.5C1 composition can be estimated from radiogenic elements[5]. If the whole Earth contains between 140 and 200 p.p.m. K, and the revised K/F ratio in C1 meteorite is taken to be 9.1 from Table 1, the content of F should be between 9×10^{19} and 13×10^{19} kg. This gives 10×10^{19} for the GA model.

Cl and Br belong to the volatiles <600 K group. Although many mineralogical details do not fit the theory of equilibrium condensation[18–20], the concentrations of volatile elements in chondritic meteorites[21] fit reasonably well with thermochemical calculations. Although similar depletion factors with respect to cosmic composition were obtained for Tl, Cl, Br and I in the bulk Earth by summation of Earth zones[3], there is an uncertainty in the concentrations of these elements. There is evidence to show that peridotite xenoliths from the mantle were metasomatized by fluids, and the bulk analyses for volatile elements used in

0028-0836/81/090762—04$01.00

Table 1 Reservoirs and chemical models

	F (kg)	P (kg)	Cl (kg)	K (kg)	Br (kg)
Hydrosphere	19×10^{14}‡	13×10^{13}‡	27×10^{18}§	57×10^{16}‡	94×10^{15}§
Sediments	—	—	18×10^{18}§	—	10×10^{15}§
Continental crust	18×10^{18}‡	29×10^{18}‡	23×10^{17}§	37×10^{19}‡	$<1\times10^{16}$§
Oceanic crust			23×10^{16}§		82×10^{14}§
Total	18×10^{18}	29×10^{18}‡	48×10^{18}	37×10^{19}‡	11×10^{16}§
GA model, whole Earth[2]	32×10^{19}	13×10^{21}	15×10^{19}	10×10^{20}	80×10^{16}
S2 model, whole Earth[1]	54×10^{19}	11×10^{21}	27×10^{19}	78×10^{19}	96×10^{16}
Sh model, whole Earth[12]	29×10^{19}	—	24×10^{19}	19×10^{20}	—
$1.5\times$Cl composition, revised*	53×10^{19}	99×10^{20}	62×10^{20}	48×10^{20}	29×10^{18}
GA model, revised†	10×10^{19}	—	12×10^{19}	—	55×10^{16}

*P and K from ref. 4 Table 2, col. *l*; F, Cl and Br from ref. 14 Table 3, row 7.
†Adjusted by ratio of abundances in ref. 13, Table 4, rows 1 and 8.
‡ From ref. 4.
§From ref. 10.

ref. 4 are higher than recent ones for selected xenoliths[22]. Nevertheless, the depletion factors for this group of volatile elements must be assumed to be constant. Ganapathy and Anders[2] modelled the group from an observed Tl/U ratio of 0.27 in terrestrial rocks and an assumed U content of 18 p.p.b. (parts per 10^9) in the whole Earth. Because the measured Tl/U ratio of 7.0 (atomic) for carbonaceous chondrites[23] is close to the value of 7.3 chosen by Cameron[13] for the cosmic composition, and the assumed U content[2] of 18 p.p.b. for the whole Earth is near the middle of the range 14–20 p.p.b. assumed here, the Tl and U content of the GA model needs no adjustment. However, revised estimates of Cl and Br are given.

Potential mineral hosts

Systematic electron microprobe analyses were made of P, F and Cl in minerals from the mantle[24,25]. Halogens were detected only in mica, amphibole and apatite. Amphibole is restricted to the outer 100 km: MARID-suite and secondary-textured crystals in peridotite xenoliths, F 0.5–1.0, Cl 0.00–0.03, K 0.3–4.3 wt%; kaersutite megacrysts in alkali basalts[26], F 0.07–0.22, Cl 0.001–0.012, K 1.0–1.7 wt%. Phlogopite-rich mica is stable down to ~200 km, with the following averages: primary texture in depleted peridotites, F 0.43, Cl 0.08, K 8.6 wt%; secondary textures of various types in peridotites, F 0.63, Cl 0.03, K 8.6 wt%. No large apatite crystals in textural equilibrium with silicate minerals of peridotite xenoliths in kimberlites have been found, but apatites from two lherzolite inclusions in basanites[27] have 0.9 and 1.0 wt% Cl. Apatites[28] from apatite-rich xenoliths in basalts from eastern Australia[29] show 2.1–2.7 wt% F and 0.17–0.22 wt% Cl. A small euhedral crystal enclosed in a dolomite bleb, itself enclosed in a Cr-diopside megacryst found in a kimberlite, contains 0.9 wt% Cl and 1.4 wt% F (ref. 30). A megacryst from the Moses Rock 'kimberlite' contains 0.5 wt% F and <100 p.p.m. Cl (ref. 25). Apatites from seven carbonatites[25] contain high F (2.0–3.7 wt%) and low Cl (0.00–0.1 wt%), and all but one apatite grain from serpentine veinlets in peridotite xenoliths from kimberlites contain low Cl. All apatites contain ~40 wt% P_2O_5, and all silicate minerals contain little P_2O_5, mostly below 0.03 wt% (ref. 24).

Complete incorporation of P into melt was adopted for a model of basalt petrogenesis from melting of mantle pyrolite[31]. A strong correlation between Ce and P in a wide range of basaltic rocks[31,32] can be explained either by the presence of most of the Ce and P in apatite which was totally consumed during melting, or by similar bulk solid–liquid partition coefficients for Ce and P (ref. 32). The good correlation between F and P in oceanic basalts and gabbros suggests that these two elements were present mainly in apatite[25], and rules out the occurrence of whitlockite unless it survived the generation of basalt. Total consumption of apatite is consistent with experiments on apatite saturation in magmas at high pressures[33].

Analyses of F, P, Cl and K in pillow interiors of tholeiitic basalts from the Mid-Atlantic Ridge[34], alkali basalts from the Azores Islands[34] and tholeiitic basalt from the East Scotia Sea back-arc basin[35] were used (Table 2) to determine whether these elements were derived solely from total incorporation of F, Cl-bearing mica and apatite into melt. The atomic ratio (F+Cl)/P for hydroxyl-free apatite is one-third. As all rocks in Table 2 have an average ratio higher than one-third, some halogens must be derived from mica. Table 2*b* shows the amounts of F and Cl needed if K were attributed to mica with the mean composition of the primary-textured type in depleted peridotite xenoliths transported by kimberlites. For the normal mid-ocean ridge basalts and the back-arc-basin basalts, the remaining F and Cl yield (F+Cl)/P ratios close to one-third. For the alkali basalts, the ratio drops to 0.08, which might be explained by: hydroxyl in apatite; lower halogen content of the mica; and derivation of K from halogen-free clinopyroxene[36]. The final option can be ruled out because the K content is too high (12,000 p.p.m.) with respect to the Ca content (~6 wt%). Only the basalts from the Azores transect have a (F+Cl)/P ratio too high to be attributed to hydroxyl-free apatite and primary-type mica, and there are at least three explanations: Cl-metasomatism by seawater; higher halogen content of mica; and precipitation of a mineral with high K/(F+Cl) in a sub-surface magma chamber.

It is assumed here that (1) for tholeiitic basalts from normal ridge segments, assignment of one-fifth of the K to clinopyroxene and four-fifths to primary-type mica would allow the apatite to contain 0.87 atoms of F and 0.13 atoms of Cl; (2) for the back-arc-basin basalts, the apatite contains about 0.67 atoms of F and 0.33 atoms of Cl; and (3) for the alkali basalts, the apatite contains about 0.75 atom of F and 0.25 atom of Cl, while the source mica has ~0.15 wt% F.

Complex processes should be expected however, during transport of a magma to the surface. Tholeiitic basalts at mid-ocean ridges are probably the residue of crystal–liquid processes in near-surface magma chambers[37], and crystallization of apatite and amphibole may cause complications. Apatites from altered gabbros and troctolites from the Mid-Cayman Rise contain 0.15–0.3 wt% Cl (no analysis for F), and most amphiboles contain little Cl (~0.01 wt%) while a few contain substantial amounts (up to 0.9 wt%)[38]. Because F enters crystals in preference to melt, surviving basaltic melt should lose F with respect to parent melt. It will now be assumed that F, Cl-rich apatite and F, Cl-bearing mica are the principal hosts for halogens and phosphorus in the upper mantle. This is consistent both with mass-balance calculations for two lherzolite xenoliths[24] which indicate that silicate minerals provide only about one-tenth of the phosphorus, and with the expectation that H_2O should fractionate preferentially into magma over apatite during crystal–liquid differentiation of the Earth, and the low abundance of primary water in mid-ocean-ridge tholeiitic basalts[39].

Table 2 Interpretation of F, P, Cl and K analyses of pillow interiors of oceanic basalts (*A–C*) and glassy pillow rims of back-arc-basin basalts (*D*)

a Bulk analyses and atomic ratios

	Weight: mean and 1σ				Atomic ratios		
Type	F (p.p.m.)	P (p.p.m.)	Cl (p.p.m.)	K (p.p.m.)	C/P	Cl/P	(F+Cl)/P
A[34]	389±131	1,110±389	454±221	5,230±2,302	0.571	0.357	0.928
B[34]	144±28	605±156	38±14	940±470	0.388	0.055	0.443
C[34]	586±118	2,441±517	350±210	12,000±3,900	0.391	0.125	0.516
D[35]	252	794	124	2922	0.517	0.136	0.653

b Calculated values assuming primary mica in source region

Type	F(mica)	Cl(mica)	F(residual)	Cl(residual)	F/P(res.)	Cl/P(res.)	(F+Cl)/P
A	261	49	128	405	0.188	0.319	0.507
B	47	9	97	29	0.261	0.042	0.303
C	600	112	−14	238	−0.009	0.085	(0.085)
D	146	27	106	97	0.217	0.106	0.323

A 20 Mid-Atlantic Ridge tholeiitic basalts from Azores transect (37–40°N).
B 16 Mid-Atlantic Ridge tholeiitic basalts from two normal ridge segments 28–33 and 49–52°N.
C 18 alkali basalts from the Azores Islands.
D 5 back-arc-basin basalts from East Scotia Sea.

Because Cl/Br ratios range only from 222 to 494 for rocks from oceanic ridges and islands[10,40], Br must have the same mineral hosts as Cl.

Tholeiitic basalts from mid-ocean ridges are more abundant that alkali basalts from oceanic islands and the diverse range of magmas in continents, and perhaps give the best clue to the average composition of the upper mantle. Although the compositions of the mineral hosts for halogens and phosphorus cannot be determined, it is suggested that the source regions of basalts contain F, Cl-rich apatite. Values of 3 wt% F and 1 wt% Cl are used for the initial calculations. Because electron microprobe analyses of Cl in apatite from upper-mantle xenoliths range from 0.2 to 1.0 wt% the effect of reducing the Cl content to 0.3 wt% will be considered. For mica, rounded values of 0.5 and 0.05 wt% will be assumed for F and Cl, respectively.

Estimate of P distribution

Subtraction of 18×10^{18} kg F and 37×10^{19} kg K for the crust from whole-Earth estimates in the revised GA model of 10×10^{19} kg F and 10×10^{20} kg K leaves 8×10^{19} kg F and 6×10^{20} kg K for the mantle. Assignment of all K to mica with 8.6 wt% K and 0.5 wt% F would require 35×10^{18} kg F. For apatite with 3 wt% F and 18 wt% P, the remaining 45×10^{18} kg F would require 27×10^{19} P. This is a small fraction of the estimated 13×10^{21} kg P in the whole Earth and hence details of mineral compositions can be ignored. Whitlockite, $Ca_3(PO_4)_2$, was ruled out as a mineral host for P in the upper mantle. The crystal structure of apatite will be assumed to be stable to high pressures. Hence it is proposed that the remaining P goes either into silicate minerals in the deep-mantle or into the core. Because the mass of the entire mantle is 40×10^{23} kg, a mean concentration of 0.3 wt% P would be required for the silicates. Although this is >10 times greater than the concentrations found in olivine, pyroxene and garnet minerals of peridotite xenoliths[24], this possibility should not be ruled out. Experimental syntheses on garnet[43] should be extended to higher pressure. The P could be assigned to the core, for which a concentration of 0.7 wt% would be required. This value lies inside the range for iron meteorites[41] (type I 0.2 wt%; type II 0.12–1.0 wt%; type III 0.07–1.2 wt%), and is consistent with Earth models based on analogy with meteorites. Assigning 0.03 wt% P to the average mantle would lower the estimate of 0.7 wt% in the core to 0.6 wt%. High-pressure reduction of phosphate to phosphide during Earth accretion should also be considered. The inferred content of 0.7 wt% P in the Earth's core is much less than the ~10 wt% S (ref. 42) required to lower the density of the outer core from that for liquid Fe, Ni and there seems to be no geophysical test of the proposed content of P. Proposed contents of C and N in the core[4] are less than for P, but also amount to a substantial fraction of whole-Earth estimates.

Estimate of Cl and Br distribution

Subtraction of the estimate of K in the crust (37×10^{19} kg) from the estimated value for the whole Earth (10×10^{20} kg) leaves 6×10^{20} kg for subcrustal regions. Assigning all this K into primary-textured mica of depleted peridotite would produce 5×10^{18} kg Cl and into secondary-textured mica 2×10^{18} kg. If the mass of F in apatite of the Earth's mantle is 45×10^{18} kg, and the apatite contains 3 wt% F and 1 wt% Cl, the mass of Cl is 15×10^{18} kg. Reduction of the Cl content of apatite to 0.3 wt% drops this mass to 5×10^{18} kg. Hence apatite seems to be a more important store of Cl in the mantle than mica, and especially than amphibole. The estimated mass of $17\text{–}20\times10^{18}$ kg Cl in the mantle (for apatite with 1 wt% Cl) implies that 25–30% of the Earth's chlorine remains in the mantle. The estimated mass of $7\text{–}10\times10^{18}$ kg Cl in the mantle (for apatite with 0.3 wt% Cl) implies that only one-seventh of the Cl remains in the mantle. The $65\text{–}68\times10^{18}$ kg and $55\text{–}58\times10^{18}$ kg are about half of the estimated whole-Earth content of 12×10^{19} kg in the revised GA model.

For Br in the mantle, a Cl/Br ratio of 300, taken from the range of 222–494 in oceanic basalts[10,40], leads to estimates of 6×10^{16} and 3×10^{16} kg Br, depending on the assumed Cl content of apatite. These values are about one-half and one-quarter of the estimate of 11×10^{16} kg in the crust and hydrosphere[10], and imply that about one-quarter to one-third of the Earth's Br remains in the mantle. The $14\text{–}17\times10^{16}$ kg is about one-third of the estimated whole-Earth value of 55×10^{16} kg in the revised GA model.

The above conclusions assume that halogens are absent from the principal minerals of the Earth's mantle. To bring the $65\text{–}68\times10^{18}$ kg Cl up to 12×10^{19} kg Cl in a revised GA model would require $\sim13\times10^{-6}$ weight fraction throughout the entire 40×10^{23} Kg of the mantle; from general crystal-chemical considerations this does not seem plausible.

Conclusions

Within the many uncertainties, I suggest (1) apatite with ~3 wt% F, up to 1 wt% Cl and 0.003 wt% Br is the principal mineral reservoir for halogens, and mica is a subsidiary reservoir; (2) apatite with ~18 wt% P is the principal store of P in the upper mantle and perhaps the lower mantle, but accounts for only one-twentieth of P in the Earth; (3) the remaining P is in a reservoir inaccessible to magmatism, and may amount to a maximum of 0.7 wt% in the core; (4) at least two-thirds to three-quarters of the Cl and Br have entered the crust, but only one-fifth of the F (Table 1) and 0.2% of the P; (5) the whole-Earth contents of halogens, K and P are consistent with less efficient capture of volatile elements than refractory ones from the solar nebula; and (6) the Cl and Br contents should be reduced in the GA model. The distribution of halogens is relevant to crystal–liquid differentiation of the Earth. If all Cl in a subducted slab is returned to the surface by island-arc volcanism[38,44,45], about one-third of the Cl in the mantle has not been transported to the surface, at least since the onset of subduction in the era of plate tectonics. If Cl was returned from the crust to the mantle during the turbulent conditions of the pre-Archaean era[7], it would not be necessary to argue that one-third of the Earth did not melt during accretion. Because present data on phase equilibria indicate that both apatite[33] and hydroxylated silicates[46] are consumed early by partial melting, it would be difficult to justify a model in which halogens could be subducted in preference to K and Rb. Simple two- or three-reservoir models[5,6] for radiogenic elements suggest that the Earth's crust was extracted from only one-half to one-third of the mantle, assuming that the original mantle was chemically uniform (see also ref. 47). This assumption of chemical uniformity may be incorrect, especially for the halogens. It is difficult to see how apatite could have been distributed uniformly throughout the mantle during accretion in view of the large amounts of accretional and core-formation heats. The possibility that apatite in the mantle tended with time to become more refractory by preferential loss of Cl, Br and OH to magma should be considered. Therefore considerable study of geochemical models for the origin and differentiation of the Earth, and further measurements of trace elements in mantle minerals are needed.

I thank A. T. Anderson, R. N. Clayton, J. B. Dawson, J. R. Goldsmith, L. Grossman, R. L. Hervig, G. Reed and especially D. M. Shaw for discussion and review, and I. Baltuska for technical help. Financial support came from grants NSF EAR77-27100 and NASA 14-001-171.

Received 9 July; accepted 9 December 1980.

1. Smith, J. V. *Miner. Mag.* **43,** 1–89 (1979).
2. Ganapathy, R. W. & Anders, E. *Proc. 5th lunar Sci. Conf.* 1181–1206 (1974).
3. Ringwood, A. E. *Origin of the Earth and Moon* (Springer, New York, 1979).
4. Smith, J. V. *Proc. 8th lunar Sci. Conf.* 333–369 (1977).
5. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84,** 6091–6101 (1979).
6. Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84,** 7411–7427 (1979).
7. Smith, J. V. *Phil. Trans. R. Soc.* (in the press).
8. O'Nions, R. K., Evensen, N. M., Hamilton, P. J. & Carter, S. R. *Phil. Trans. R. Soc.* A**288,** 547–559 (1978).
9. Sun, S. S. & Nesbitt, R. W. *Earth planet Sci. Lett.* **35,** 429–448 (1977).
10. Schilling, J.-G., Unni, C. K. & Bender, M. L. *Nature* **273,** 631–636 (1978).
11. Taylor, S. R. *Proc. 10th lunar Planet. Sci. Conf.* 2017–2030 (1979).
12. Shaw, D. M. *Precamb. Res.* **10,** 281–299 (1980).
13. Cameron, A. G. W. *Space Sci. Rev.* **15,** 121–146 (1973).
14. Dreibus, G., Spettel, B. & Wänke, H. *Phys. Chem. Earth* **11,** 33–38 (1979).
15. Van Schmus, W. R. & Ribbe, P. H. *Geochim. cosmochim. Acta* **33,** 637–670 (1969).
16. Goldberg, R. H., Burnett, D. S., Furst, M. J. & Trombrello, T. A. *Meteoritics* **9,** 347 (1974).
17. Allen, R. O., Jr. & Clark, P. J. *Geochim. cosmochim. Acta* **41,** 581–586 (1977).
18. Hutchison, R., Bevan, A. W. R., Agrell, S. O. & Ashworth, J. R. *Nature* **280,** 116–119 (1979).
19. Bunch, T. E. & Chang, S. *Geochim. cosmochim. Acta* **44,** 1543–1577 (1980).
20. Leitch, C. A. & Smith, J. V. *Nature* (in the press).
21. Grossman, L. & Larimer, J. W. *Rev. Geophys. Space. Phys.* **12,** 71–101 (1979).
22. Jagoutz, E. *et al. Proc. 10th lunar Planet. Sci. Conf.* 2031–2050 (1979).
23. Krähenbühl, U., Morgan, J. W., Ganapathy, R. & Anders, E. *Geochim. cosmochim. Acta* **37,** 1353–1370 (1973).
24. Bishop, F. C., Smith, J. V. & Dawson, J. B. *Lithos* **11,** 155–173 (1978).
25. Smith, J. V., Delaney, J. S., Hervig, R. L. & Dawson, J. B. *Lithos* (in the press).
26. Garcia, M. O., Muenow, D. W. & Liu, N. W. K. *Am. Miner.* **65,** 306–312 (1980).
27. Frey, F. A. & Green, D. H. *Geochim. cosmochim. Acta* **38** 1023–1060 (1974).
28. Smith, J. V., Hervig, R. L., Wass, S. Y. & Dawson, J. B. (in preparation).
29. Wass, S. in *Inclusions in Kimberlites and other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A. 366–373 (American Geophysical Union, Washington, 1979).
30. Hervig, R. L. & Smith, J. V. *Amer. Miner.* (in the press).
31. Sun, S. S. & Hanson, G. N. *Geology* **3,** 297–302 (1975).
32. Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19,** 463–513 (1978).
33. Watson, E. B. *Geophys. Res. Lett.* **6,** 937–940 (1979).
34. Schilling, J.-G., Bergeron, M. B. & Evans, R. *Phil. Trans. R. Soc.* A (in the press).
35. Muenow, D. M. Liu, N. W. K., Garcia, M. O. & Saunders, A. D. *Earth planet. Sci. Lett.* **47,** 272–278 (1980).
36. Smith, J. V., Hervig, R. L., Ackermand, D. & Dawson, J. B. in *Kimberlites, Diatremes and Diamonds* (eds Boyd, F. R. & Meyer, H. O. A.) (American Geophysical Union, Washington, 1979).
37. Bryan, W. B. & Moore, J. G. *Bull. geol. Soc. Am.* **88,** 556–570 (1977).
38. Ito, E. thesis, Univ. Chicago (1979).
39. Delaney, J. R., Muenow, D. W. & Graham, D. G. *Geochim cosmochim. Acta* **42,** 581–594 (1978).
40. Unni, C. K. & Schilling, J.-G. *EOS* **61,** 407 (1980).
41. Buchwald, V. F. *Handbook of Iron Meteorites* (University of California, Berkeley, 1975).
42. Ahrens, T. J. *J. geophys. Res.* **84,** 985–998 (1979).
43. Thompson, R. N. *Earth planet. Sci. Lett.* **26,** 417–424 (1980).
44. Anderson, A. T. *Bull. Am. geol. Soc.* **85,** 1485–1492 (1979).
45. Anderson, A. T. *Rev. Geophys. Space Phys.* **13,** 37–55 (1975).
46. Wyllie, P. J. *Am. Miner.* **64,** 464–500 (1979).
47. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).

Population dynamics of fox rabies in Europe

Roy M. Anderson[*], Helen C. Jackson[†], Robert M. May[‡] & Anthony M. Smith[*]
[*] Zoology Department and Centre for Environmental Technology, Imperial College, London University, London SW7 2AZ, UK
[†] WHO Collaborating Centre for Rabies Surveillance and Research, Federal Research Institute for Animal Virus Diseases, D7400 Tübingen, Postfach 1149, FRG
[‡] Biology Department, Princeton University, Princeton, New Jersey 08450, USA

A simple mathematical model for the overall dynamics of the interaction between fox populations and rabies is presented. The model helps to explain epidemiological patterns observed in Europe, including the 3 to 5 year cycle in fox populations infected with rabies, threshold densities and average levels of prevalence of infection. We give a quantitative discussion of the possibilities of controlling rabies by culling or vaccinating foxes (or by a mixture of the two).

RABIES is a directly transmitted viral infection of the central nervous system, to which all mammals are thought to be susceptible (although not equally so)[1-4]. Throughout human history it has been one of the most feared of all infectious diseases[1,5-9]. Today, the number of people dying from rabies in Europe is insignificant (around 4 per annum in recent years), and even the global total (estimated as high as 15,000 per annum, mainly in India)[10] is dwarfed by other diseases. However, the nature of the clinical symptoms, which are distressing and often hideous, plus the knowledge that death is virtually inevitable once symptoms appear[8,11], makes rabies a continuing cause for concern.

Many major epidemics of rabies have been recorded in Europe over the past seven centuries[1,8], with the dog acting as the main host and primary transmitter of the infection to man. The current epidemic in Central Europe is thought to have originated in Poland in 1939 (ref. 12), and is characterized by a high incidence in the red fox *Vulpes vulpes*; of the 16,820 cases of animal rabies reported in Europe in 1979, over 70% were in this host species[13]. Throughout Europe, and to a lesser extent in Canada and the United States, foxes (red, grey and arctic species) play a dominant role in the transmission and maintenance of the disease, although other mammals (including dogs, skunks, badgers, racoons, wolves and bats) are important in these and other regions of the world[1,7,14-17].

The current European epidemic has spread outwards from Poland at an annual rate of 30–60 km, with the present front in France being less than 40 km from the English Channel coastline (Fig. 1). The introduction of rabies into Britain through illegally imported pets[18] therefore seems a distinct possibility and is especially worrying because of the uniquely high urban and suburban density of foxes, side by side with dense populations of dogs and cats[19,20].

0028-0836/81/090765—07$01.00

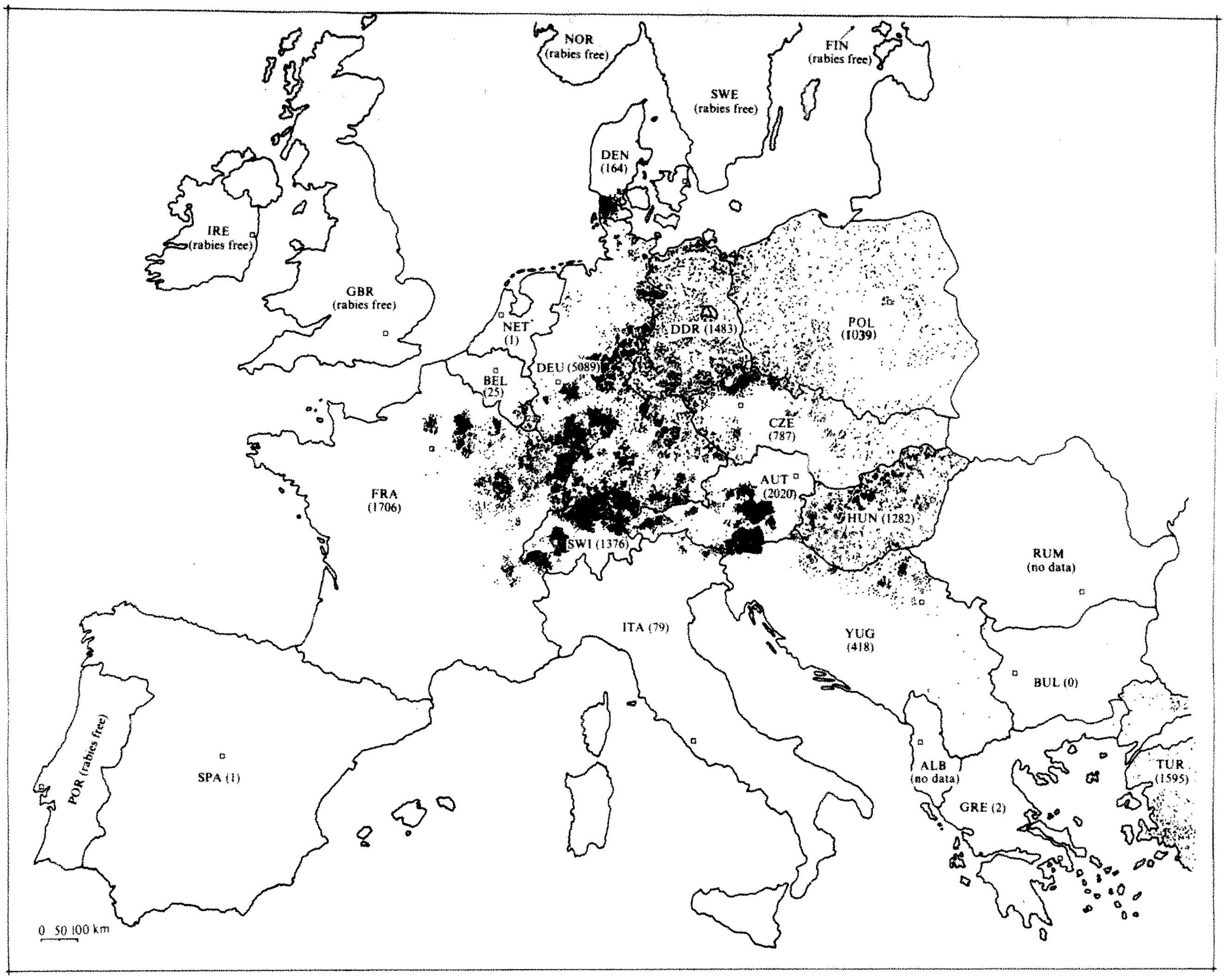


Fig. 1 Geographical distribution of reported European cases of rabies, both in wildlife and domestic animals, during 1979. Each dot denotes one reported case, excepting a few dots in northern Switzerland and southern Germany which represent between one and four cases.

Previous studies of the epidemiology of rabies in fox populations have focused on the characteristics of the viral infection within the host[1,14,21–26], on the ecology of the fox[14,18–20,27–33,101] and on computer simulations along with statistical studies of the spatial and temporal development of epidemics[34–43]. This article, however, takes a different approach in studying the overall population biology of the interaction between foxes and rabies with the emphasis on recurrent epidemic behaviour (in contrast to the more usually studied spatial aspects of rabies epidemiology). We present a simple mathematical model which exhibits patterns of dynamical behaviour in striking agreement with observed epidemiological facts. The model enables us to identify the basic factors controlling persistence of the pathogen, and to assess the effectiveness of different methods for controlling the disease.

Basic model and epidemiological parameters

We use a deterministic, compartmental model in which the fox population is divided into susceptibles, infecteds that are not yet infectious, and infectious individuals; these three classes are defined to have population densities X, I and Y, respectively[44,45]. The model has no category of recovered, immune foxes, because few, if any, foxes recover once the virus is established in the host[1–3,7,8,24,27]. The total density of foxes, N, is thus $N = X + I + Y$.

Foxes are territorial,a family's territory size depending on the type of habitat, food availability and the density of foxes[19,20]. Home ranges in Europe and North America typically lie in the range 2.5 to 16 km^2 (refs 46–52). In the absence of rabies, fox populations seem to increase up to some characteristic density, K, determined by the carrying capacity of the habitat. The value of K is set by the mixture of habitat types in a defined area[19]; within Europe, K is typically in the range 0.1–4 foxes per km^2 (refs 19, 20, 28, 52–54) although much higher densities have been reported in suburban and urban areas of Britain[19,20].

As the death rate of young foxes is density dependent[19,20,28–32], we assume the *per capita* death rate is linearly related to N, such that the net rate is $(b + \gamma N)N$; here, $1/b$ denotes fox life expectancy in the absence of resource limitation, and γ measures the severity of density-dependent constraints. Life expectancy, $1/b$, is typically in the range 1.5–2.7 yr within Europe (Table 1*a*); we take $1/b$ to be 2 yr (Table 1*b*).

Wandeler[55] argues that the rate of reproduction is independant of population density for foxes in Europe, although other reports suggest it may be limited by food availability[30,31,53,54,101]. For simplicity, we assume that density dependence acts primarily on mortality, and that the *per capita* birth rate is a constant, a. The number of cubs in a litter ranges from 1 to 10, with a mean in Europe of 4.7 (refs 14, 18–20, 52, 55). Sex ratios are in general close to unity at birth, and the pregnancy rate is in the region of 90% (refs 19, 20, 55, 56) with a further 10% of vixens failing to produce offspring. By combining these values we obtain for a an average value of 1.0 per annum. The intrinsic *per capita* population growth rate, $r = a - b$, is then $\simeq 0.5$ per annum.

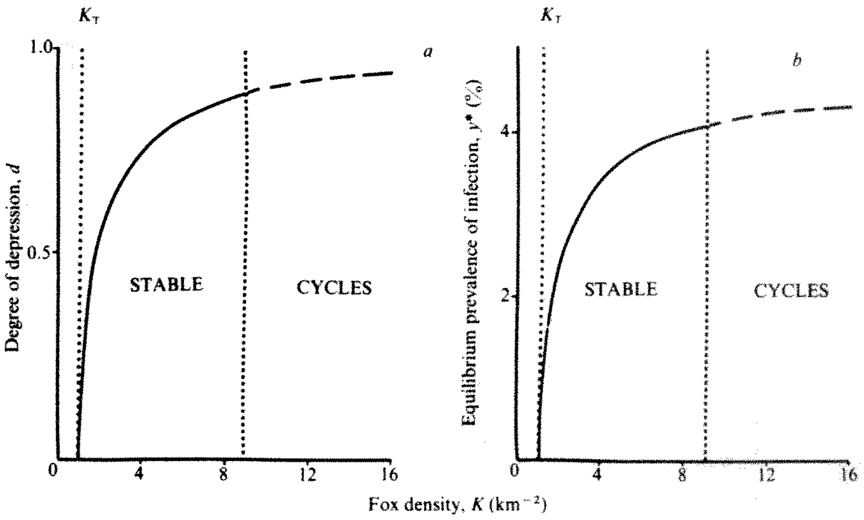


Fig. 2 *a*, Endemic rabies suppresses fox population density to a disease controlled equilibrium N^* where $N^* = K[(\sigma + a)(\alpha + a) - ar]/(\sigma\beta K - ar)$. This steady state density is necessarily less than the disease-free density K, and the degree of population depression induced by rabies, defined as $d = 1 - N^*/K$, is plotted for various values of K. In the region denoted 'stable' the model exhibits damped oscillations to a stable state, and in the region labelled 'cycles' limit cycle behaviour occurs and the dashed line denotes the average degree of depression throughout one cycle. Parameter values again are as listed in Table 1*b*. *b*, This graph portrays the relationship between the equilibrium prevalence of infection y^* and K, where $y^* = (Y^* + I^*)/N^*$, with Y^* and I^* the equilibrium densities of infectious and infected but not yet infectious (latent) foxes. By setting $\mathrm{d}X/\mathrm{d}t = \mathrm{d}I/\mathrm{d}t = \mathrm{d}Y/\mathrm{d}t = 0$ in equations (2)–(4) in the text we obtain $y^* = [(\alpha + \sigma + a - v)/(\sigma + a)][r/(\alpha + a)]D$ where $r = r(1 - N^*/K)$, $D = [1 - 1/R]/[1 - (ar)/[(a + \alpha)(a + \sigma)]]$ and R is as defined in equation (5) of the text. The regions labelled 'stable' and 'cycles' are as defined in the legend to graph *a* and the dashed line in the cycles region denotes the average prevalence of rabies throughout one cycle. Note that both the degree of depression, *d* (graph *a*) and the prevalence y^* (graph *b*) rise as K increases.

These assumptions about the biology of the fox population lead to the familiar logistic equation[57] for the dynamics, in the absence of rabies:

$$\mathrm{d}N/\mathrm{d}t = rN(1 - N/K) \tag{1}$$

Here, the carrying capacity of the habitat at equilibrium is $K = r/\gamma$ (in practice, we try to estimate K itself, rather than γ).

Following conventional lines[58], we assume that the rate at which foxes acquire rabies is proportional to the number of encounters between susceptible and infectious foxes, βXY. Here, β is a transmission coefficient, with $1/\beta$ being proportional to the average time interval between fox contacts; reports suggest that individuals from adjacent family groups meet, on average, every 4 to 6 days[20]. Foxes are assumed to pass from the latent or incubating class to the infectious state at a *per capita* rate σ, such that the average incubation period is $1/\sigma$. This period is variable for foxes (from 12 to 110 days) and depends on factors like age, site of infection and nutritional state[59]; on average, $1/\sigma$ is around 28–30 days[21–24]. Infectious or 'rabid' foxes have a short life expectancy, varying from 3 to 10 days, with an average of 5 days[14,21–25]. We assume rabid animals die at a constant *per capita* rate α; thus, $1/\alpha$ is their life expectancy.

Table 1 Life expectancy ($1/b$) of the red fox and values of epidemiological and other parameters

a	Life expectancy (yr)	Study area
	1.5	Denmark[97]
	1.7	Wales[18]
	1.8	Scotland[28]
	1.9	Suburban London[98]
	2.2	Germany[99]
	2.7	Netherlands[100]

b	Symbol	Meaning	Value
	b	Average *per capita* intrinsic death rate	0.5 yr^{-1}
	a	Average *per capita* birth rate of foxes	1 yr^{-1}
	r	*Per capita* population growth rate, $a - b$	0.5 yr^{-1}
	K	Fox carrying capacity, $K = r/\gamma$	Various
	σ	$1/\sigma$ is the average latent period	13 yr^{-1}
	α	Death rate of 'rabid' foxes	73 yr^{-1}
	β	Transmission coefficient (estimated using equation (7)) for threshold density, $K_T = 1.0$ foxes per km^2	79.69 km^2 yr^{-1}

Because the infection is severely debilatating and of relatively short duration, we assume that infected foxes do not contribute to the reproductive effort of the population.

These biological assumptions lead to the following set of first order differential equations for the dynamics of fox populations infected with rabies:

$$\mathrm{d}X/\mathrm{d}t = rX - \gamma XN - \beta XY \tag{2}$$

$$\mathrm{d}I/\mathrm{d}t = \beta XY - (\sigma + b + \gamma N)I \tag{3}$$

$$\mathrm{d}Y/\mathrm{d}t = \sigma I - (\alpha + b + \gamma N)Y \tag{4}$$

The equation for the total population is obtained by adding the above three:

$$\mathrm{d}N/\mathrm{d}t = aX - (b + \gamma N)N - \alpha Y \tag{5}$$

Table 1*b* summarizes our estimates of the values for the relevant parameters.

Population dynamics

The above formulation differs from conventional epidemiological models[58] in that the host population N is a dynamical variable, affected both by the disease and by density-dependent constraints imposed by resource limitation. Rabies will be maintained within the fox population provided the 'basic reproductive rate', R, of the infection exceeds unity[60,61]; R is the expected number of secondary cases produced during the lifespan of an infectious fox introduced into a population of K susceptible animals. For the system defined by equations (2)–(4),

$$R = \frac{\sigma\beta K}{(\sigma + a)(\alpha + a)} \tag{6}$$

The condition $R > 1$ for endemic maintenance of rabies may equivalently be expressed as the requirement that the fox population exceed a 'threshold density[58,62], $K > K_T$, with the definition

$$K_T = (\sigma + a)(\alpha + a)/\beta\sigma \tag{7}$$

If $R < 1$, or equivalently $K < K_T$, the fox population will settle to its disease-free equilibrium density K.

Epidemiological evidence suggests K_T lies in the range 0.25–1.0 foxes per km^2, with the most frequently reported value in Europe being around 1.0 foxes per km^2 (refs 18, 63–65). This observation can now be combined with equation (7), along with the parameter values listed in Table 1*b*, to give a rough estimate of the transmission parameter β. The value $\beta \approx 80$ yr (per unit density of foxes) thus obtained agrees remarkably well with independently obtained behavioural observations in rabies-free

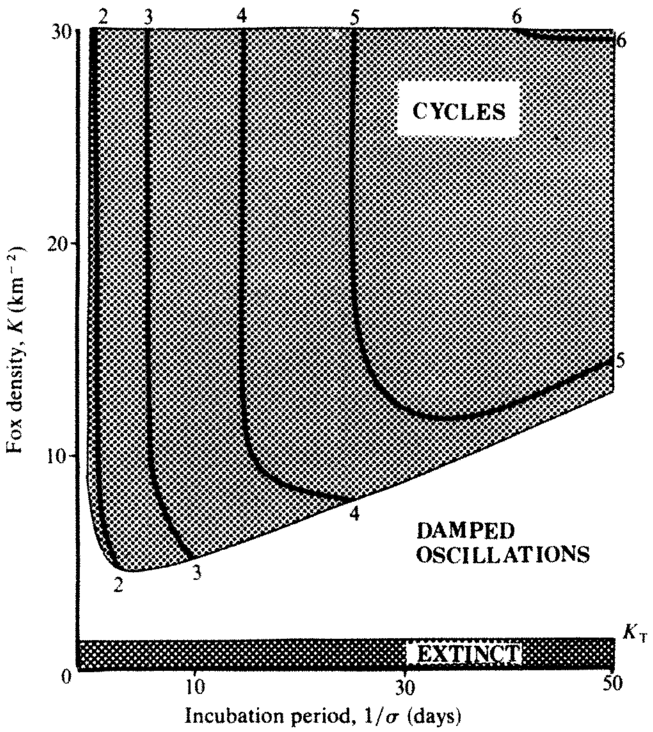


Fig. 3 The three regimes of dynamical behaviour that can be exhibited by the model defined in equations (2) to (4) are illustrated here in the $K - 1/\sigma$ parameter space. As discussed in the text these three regimes are: rabies unable to maintain itself within the fox population ($R < 1$), labelled 'extinct'; fox density controlled to stable equilibrium, N^*, by disease ($R > 1$), labelled 'damped oscillations; and limit cycle behaviour in fox abundance and disease prevalence ($R > 1$), labelled 'cycles'. In this limit cycle region, the contour lines correspond to specific periods of oscillation, labelled in years. All rate parameters are expressed in units of yr^{-1}, and have the values listed in Table 1*b*.

populations concerning the average time interval between fox contacts in regions where the average density is around 1 fox per km^2 (which gives a value for $1/\beta$ of 4 or 5 days, or $\beta \simeq 70$–80^{-1} at this density)[19]. Such agreement seems surprising in view of the belief that rabid foxes have higher contact rates than do healthy animals, although altered host behaviour as a consequence of infection may act to compensate for the fact that only a certain proportion of contacts between healthy and rabid animals will result in disease transmission. When $R > 1$, rabies will regulate fox abundance below the disease-free level K (Fig. 2*a*). This regulated state, however, may be a stable constant value, or it may be a stable cycle; the cyclic solutions tend to arise when the carrying capacity of the fox habitat is relatively large (K significantly larger than K_T).

Figure 3 shows the types of dynamical behaviour which can occur, as functions of the latency interval $1/\sigma$ and the density of the disease-free fox population K, with the other epidemiological parameters having the values listed in Table 1*b*. The values of σ and K discussed above can be seen to suggest cycles with periods in the general range of 3–5 yr. Such periods are precisely what are observed in Europe and North America, where a striking feature of rabies epidemiology is the regular 3-, 4- or 5-yr oscillations in fox density and in the prevalence of infection[6,14,18,24,27,66-69]. The cycles are most commonly observed in regions where fox densities were high before the introduction of rabies, and cyclic behaviour is often absent in low density populations[15]; these tendencies agree with the predictions of the model, as shown in Fig. 3. Qualitative insight into these properties of our model may be obtained by noting that the period of the cycle is determined primarily by the intrinsic growth rate, r, of the fox population (although, as Fig. 3 shows, other factors also influence the period), with rabies acting as a time-delayed density-dependent regulator of fox population growth (the length of the time lag being determined by how long the fox density is below K_T); it is well understood that such mechanisms can induce oscillatory behaviour[70].

For the parameter values most relevant to the epidemiology of rabies in Europe ($1/\sigma \simeq 30$ days, K in the region 1–4 foxes per km^2), Fig. 3 shows the model to predict damped oscillations in fox density and disease prevalence, with periods around 3–4 yr, leading eventually to a stable disease-controlled equilibrium. As illustrated in Fig. 4, however, the damping time is long, being of the order of 40 yr or more from the first appearance of the disease. This propensity to prolonged oscillations, even in the dynamical regime labelled 'damped oscillations' (Fig. 4), is further accentuated by seasonal changes in the values of pertinent parameters. Fox reproduction is clearly seasonal[19,20], and in addition the behavioural ecology of the fox (particularly mating and dispersal patterns) imposes a seasonal trend on disease transmission, which generates a marked peak in disease prevalence from January to April[14,15,24,27]. Numerical studies show that seasonality in the transmission parameter β accentuates the non-seasonal oscillations in prevalence and fox density, considerably lengthening, for example, the damping time in the system portrayed in Fig. 4[71]. In short, the model gives a fairly robust prediction of 3–5-yr cycles in European fox populations infected with rabies[101].

Before leaving the subject of cycles, one further complication deserves mention. In Northern Europe, 4-yr cycles have been reported for rabies-free fox populations; these cycles are likely to be a consequence of the well known 4-yr cycles in populations of microtine rodents in northerly regions[27,28,72-75]. We are, however, not aware of any such cycles in rabies-free fox populations in Central or Southern Europe.

As a direct consequence of its high pathogenicity, rabies will only persist within fox populations at very low levels of prevalence; this is an example of the commonly observed biological relationship between low 'standing crop' and high rate of turnover. Persistence is in part ensured by the high susceptibility of foxes to rabies. In endemic areas of stable rabies, prevalence seems to lie in the range of 3–7% (ref. 76), although significant fluctuations are caused by seasonal effects and by the intrinsically oscillatory nature of the interaction between host and pathogen. As illustrated in Fig. 2*b*, our model gives prevalence values which agree with those observed. Moreover, the model predicts that the prevalence of rabies will be higher in favourable fox habitats (large K), which is consistent with the epidemiological evidence in Germany[14,27,77].

During the troughs in fox abundance (see Fig. 4) the prevalence may drop to very low levels. In these circumstances, the chance of stochastic effects causing extinction of the disease is high, particularly in small areas where the fox habitat is good (and consequently the amplitude of the cycles is large). The persistence of disease over large areas is probably ensured by the high spatial mobility of the fox host; young males may disperse

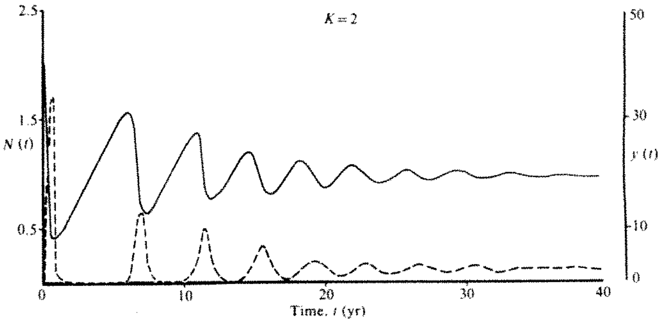


Fig. 4 A numerical solution of the model defined in equations (2) to (4) using the parameter values of Table 1*b*. The carrying capacity of the fox habitat, K, is set at 2 animals per km^2 and $K_T = 1$ fox per km^2. The solid line represents fox density at time t, $N(t)$, and the dashed line the prevalence of infection, $y(t)$.

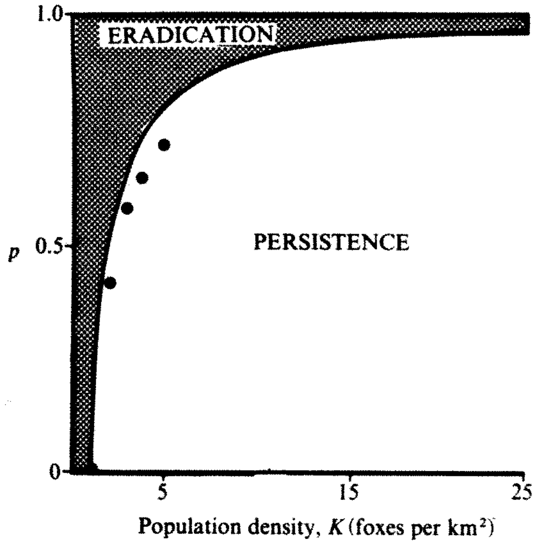


Fig. 5 This figure depicts the relationship between the proportion of the fox population p, protected by vaccination, and the disease-free carrying capacity of the fox habitat K; the threshold density is taken to be $K_T = 1.0$ foxes per km^2. In the shaded region equation (11) in the text is satisfied and rabies cannot persist; in the unshaded region equation (11) is not satisfied and the disease maintains itself. The observation of equation (11) is based on an extension of the model defined in equations (2) to (4) in the text, to include an immune category of foxes of density $2(t)$ at time t; susceptible animals are vaccinated at a *per capita* rate ϕ and immune animals lose immunity at a *per capita* rate δ. The new model is $\mathrm{d}X/\mathrm{d}T = a(X+Z)-(b+\gamma N)X-\beta XY-\phi X+\delta Z$, $\mathrm{d}I/\mathrm{d}t = \beta XY-(\sigma+b+\gamma N)I$, $\mathrm{d}Y/\mathrm{d}t = \sigma I-(\alpha+b+\gamma N)Y$, $\mathrm{d}Z/\mathrm{d}t = \phi X-(\delta+b+\gamma N)Z$. For disease eradication the equilibrium solution (obtained by setting $\mathrm{d}X/\mathrm{d}t = \mathrm{d}I/\mathrm{d}t = \mathrm{d}Y/\mathrm{d}t = \mathrm{d}Z/\mathrm{d}t = 0$) of interest is the disease-free state defined by $Y^* = I^* = 0$ and $N^* = K$. This state is locally stable provided $\phi > (a+\delta)(R-1)$, and the proportion of vaccinated foxes, p, is $p = Z^*/(Z^*+X^*) = \phi/(\phi+\mathrm{a}+\delta)$ which may alternatively be expressed as equation (11) in the text. The four solid circles are the numerical results obtained by Berger[34] from a very detailed model, as discussed more fully in the text. The circles denote the boundary, predicted by Berger's model, between levels of vaccination which result in either disease persistence or eradication.

13 km or more from their birth place, and adult foxes often travel 11–13 km each night[19,20]. Empty territories created by deaths are rapidly occupied by itinerant and mobile male foxes[20]. The involvement of other host species may also be important for the maintenance of endemic rabies during periods when fox abundance is low.

Control by reducing fox density

Attempts to control the spread of rabies in Europe and North America by hunting, gassing, poisoning or trapping foxes have generally met with little success[8,15,19,24,30,76,78]; the disease is still endemic in Poland and West Germany despite intensive efforts to reduce fox density[14,27]. Theory suggests the aim of such programmes should be to maintain fox numbers below the threshold density K_T, given by equation (7). Encouraged by the agreement with observed epidemiological patterns, we now apply our model to explore this problem.

There is a large body of literature, particularly for fisheries and whaling, dealing with the dynamics of natural populations under various regimes of culling or harvesting[79–81]. One approach is to set a constant quota of animals, Λ, to be removed annually in order to keep the fox density at some specified level, below K. In the absence of rabies, the dynamics of the fox population under such culling obeys

$$\mathrm{d}N/\mathrm{d}t = rN(1-N/K)-\Lambda \qquad (8)$$

The analysis of this situation now follows standard lines laid down in the fisheries literature[79–81]. Two cases arise.

If $K_T > \frac{1}{2}K$, the threshold density lies above the point where the fox population growth curve has its maximum (the 'maximum sustainable yield', MSY, point), and stable control to a density below K_T is possible, provided that the culling rate Λ exceeds Λ_T

$$\Lambda_T = rK_T(1-K_T/K) \qquad (9)$$

For K_T in the region of 1 fox per km^2, this regime applies to areas of poor fox habitat, where K is below 2 foxes per km^2; culling can maintain the fox population at a stable equilibrium density $N^* < K_T$.

If $K_T < \frac{1}{2}K$, two complications arise. First, the initial quota must exceed the MSY value, $\Lambda > rK/4$, to drive fox densities below the point (at $\frac{1}{2}K$) where the population growth curve has its peak. Second, the fox population cannot be stably maintained at a finite value below $\frac{1}{2}K$ by culling; the equilibrium point at K_T, attained by culling at the rate Λ_T of equation (9), is unstable. In principle, the population could be driven to extinction by operating at $\Lambda > rK/4$, but in practice a prohibitively large effort would be required to meet this quota when fox density was very low. This regime applies to areas of good fox habitat where $K > 2$ foxes per km^2; such populations will be difficult to control by culling because of the continually high effort needed to keep an area free of foxes, given their ability to disperse over large areas and rapidly to recolonize vacant territories[18,19,82,83].

A different approach is required to devote some specified constant effort (for example, constant number of man-hours) to killing foxes. This regime, in effect, changes the intrinsic *per capita* death rate of foxes from b to $b+\Delta b$, where Δb is the additional mortality imposed by constant effort culling. Fox densities will be maintained below K_T, and rabies eradicated from the population, provided

$$\Delta b/r > 1-1/R \qquad (10)$$

Here, R is the 'basic reproductive rate' for rabies in the original population, as defined in equation (6). Again, we see that if R is large, as it typically is in good fox habitat, Δb must essentially equal r itself, which implies a considerable level of effort. The practical implications of these predictions need to be assessed particularly with respect to the relationship between effort (determined by r and K), fox density and the rate of mortality achieved by control measures.

Other, more elaborate, culling regimes could be considered, but the broad theoretical conclusion remains that the control of rabies by culling foxes will be difficult to achieve in all except poor habitats, and even in these continual effort must be applied. Such conclusions agree with recent experience in Europe, and are a consequence of the high reproductive potential and dispersal behaviours of foxes[18,20,82,83].

The use of pathogens such as canine distemper or canine hepatitis as suppressors of fox numbers may be a feasible alternative, provided the threshold density for maintenance of the introduced pathogen is lower than that for rabies. Field observations in North America suggest that a more detailed appraisal of this approach is warranted[66].

Control by vaccination

The possibilities of immunizing wild foxes against rabies have been vigorously pursued in recent years[84–87], even though many problems surround this approach (not the least being the danger than an attenuated live vaccine which effectively immunized foxes might introduce a mild form of the disease into other host species)[1,19,88]. Substantial progress has been made, however, in the development of oral vaccines and in techniques for delivering them in baits[88–90]. Extensive field trials have been carried out, for example, in Canada, Germany and Switzerland[19,89,91,92]. In the design of these trials, there has been much discussion about the level of protection needed within the fox population to eliminate rabies[19,34,36,43].

Our model provides an answer: the proportion of the population, p, that must be protected at any one time is simply

$$p > 1 - 1/R \qquad (11)$$

Here, R is as defined in equation (6) (see also Fig. 5 legend)[60,61,71]. This general relation (11) is illustrated in Fig. 5. Because R may equivalently be written as $R = K/K_T$, we see that K_T is the only parameter involved in specifying p as a function of K.

Figure 5 also displays the results of an elaborate numerical simulation by Berger[34]. His model involves a detailed statistical description of dispersal by, and contact between, foxes on a large two-dimensional grid, along with statistical distributions for the duration of latent and infectious states in individual foxes. Like us, Berger took K_T to be 1 fox per km^2, and there is good agreement between the results of this detailed study and the simple equation (11); Fig. 5 provides a good example of the basic understanding that can be gained from simple models.

Equation (11) and Fig. 5 show clearly that the required degree of protection is greater in dense fox populations than in sparse ones; for K around 2 foxes per km^2, roughly 50% of the population must be protected (given the K_T of Table 1*b*). Vaccination thus seems to be a better method than culling, particularly in view of the 74% level of protection achieved by Johnston in Canada[19,93] and the 60–70% bait acceptance attained in Germany[89]. In good fox habitats, however, the required protection level approaches 100% (see Fig. 5), and will be difficult to achieve.

Control by vaccination and culling

If vaccination (of a proportion p of the fox population) is combined with constant effort culling (which adds Δb to the fox death rate), the criterion for eradication of rabies becomes

$$p + q - pq > 1 - 1/R \qquad (12)$$

Here, $q = \Delta b/r$. This criterion generalizes equations (10) and (11), to provide an explicit estimate of the efficacy of a combined strategy. For instance, if $R = 4$ we would need to vaccinate 75% of the foxes, or to kill at a rate $\Delta b/r = 75\%$, under a pure strategy; a combined strategy would require, for example, $p = 50\%$ and $\Delta b/r = 50\%$. Without making any estimate of the relative costs, it does seem likely that the expenses of vaccination or of culling will rise faster than linearly with increasing p or Δb, so that the combined strategy may often be the most cost effective (as well as permitting more foxes to live than does pure culling).

Conclusion

In the event of an outbreak of rabies among wildlife in Britain, the government's official policy is to concentrate on the destruction of foxes[18,19,94]. Our theoretical results, analysed in the light of available data, support MacDonald's conclusion[19] that such an approach is unlikely to be effective once rabies becomes established within the fox population[83,92,95]. Vaccination, conditional on the development of a 'safe' vaccine[96], and supported by culling in low density fox populations, is more likely to be successful. Our model provides a quantitative framework for discussing the precise level of control (whether vaccination, or culling, or a mixture) required in different fox habitats.

In view of the predicted difficulties of controlling rabies within areas of good fox habitat, the most sensible priority would be the compulsory vaccination of dogs and cats, along with eradication of strays and feral populations, so as to minimize the risk of human exposure in urban and suburban regions of Britain.

When assessing the significance of these conclusions it is clearly important to bear in mind the limitations of simple deterministic models (based on assumptions of homogeneous mixing), which ignore spatial and stochastic effects, and the accuracy of the data base used for parameter estimation. The power of simple formulations lies in the relative ease with which relationships can be established between important variables, a facility rarely provided for by complex simulation models (see equation (12) and Fig. 5). We emphasize, however, that our model predicts patterns of population behaviour in striking agreement with observed epidemiological trends. This leads us to believe that, by the specification of a few easily measured parameters, the model captures the significant features of the population association between host and pathogen and is thus of practical value to the design and implementation of control measures. To improve the accuracy of model predictions, however, additional data are ideally required, particularly with respect to the threshold density, K_T, the demography of fox populations in Britain and the dispersal behaviour plus contact rates of rabid foxes.

This work was supported in part by the Department of the Environment (R.M.A., H.C.J. and A.M.S.) and by the NSF under grant DEB77-01565(R.M.M.). We acknowledge the WHO Collaborating Centre for Rabies Surveillance at the Federal Research Institute for Animal Virus Disease, Tübingen, for the provision of data concerning reported rabies cases in Europe and the production of the map presented in Fig. 1.

Received 20 August; accepted 27 November 1980.

1. Baer, G. M. *The Natural History of Rabies* Vols I, II (Academic, New York 1975).
2. Sikes, P. K. *Am. J. vet. Res.* **23,** 1041–1047 (1962)
3. Parker, R. L. & Wilsnack, R. E. *Am. J. vet. Res.* **27,** 33–46 (1966).
4. Constantine, D. G. *Am. J. vet. Res.* **27,** 13–32. (1966).
5. Henschen, F. *The History of Diseases* (Longman, London, 1966).
6. Nikolitsch, M. *Blue Book for the Veterinary Profession* Vol. 10, 7–12 (1965).
7. Nageno, Y. & Davenport, F. M. (eds) *Rabies* (University Park Press, Baltimore, 1971).
8. Kaplan, C. (ed.) *Rabies: The Facts* (Oxford University Press, 1977).
9. Steele, J. H. *Scand. J. infect. Dis.* **5,** 299–312 (1973).
10. Schwabe, C. W. *Rep. 1st Wld Hlth Org. Semin. on Veterinary Public Health* (Mukteswar, New Delhi, 1971).
11. Fenner, F., McAuslan, B. R., Mims, C. A., Sambrook, J. & White, P.O. *The Biology of Animal Viruses* (Academic, New York, 1974).
12. Wld Hlth Org. *Features* **1,** 1–4 (1971).
13. Wld Hlth Org. *Rabies Bull. Europe* Vol. 3, 5 (WHO Collaborating Centre for Rabies Surveillance and Research, Tübingen, 1980).
14. Wandeler, A. *et al. Zentbl. vet. med.* **B21,** 735–56, 757–64, 765–73 (1974).
15. Johnston, D. H. & Beauregard. *M. Bull. Wildl. Dis. Ass.* **5,** 357–70 (1969).
16. Steele, J. H. *Scand. J. infect . Dis.* **5,** 299–312 (1973).
17. Chalmers, A. W. & Scott, G. R. *Trop. Anim. Hlth Prod.* **1,** 33–55 (1969).
18. Lloyd, H. G. *Trans. R. Soc. trop. Med. Hyg.* **70,** 179–87 (1976).
19. MacDonald, D. W. *Rabies and Wildlife. A Biologist's Perspective* (Oxford University Press, 1980).
20. Lloyd, H. G. *The Red Fox* (Batsford, London, 1980).
21. Winkler, W. G. *J. infect. Dis.* **125,** 674–675 (1972).
22. Sikes, R. K. in *Infectious Diseases of Wild Mammals* (eds Davis, J. W., Karstad, L. H. & Trainer, D. O.) 3–19 (Iowa State University Press, Ames, 1970).
23. Tierkel, E. *Symp. Ser. Immunbiol. Stand.* **1,** 245–258 (1966).
24. Toma, B. & Andral, L. *Adv. Virus. Res.* **21,** 1–36 (1977).
25. Bisseuru, B. *Rabies* (Heinemann, London, 1972).
26. Blancou, J. Aubert, M. F. A., Andral, L. & Artois, M. *Rev. Med. Vet.* **130,** 1001–1015 (1979).
27. Kauker, E. & Zettl, K. *Dt. tierarztl. Wschr.* **67,** 463–467 (1960).
28. Klob, H. H. & Hewson, R. *J. appl. Ecol.* **17,** 7–19 (1980).
29. Harris, S. *Mammal. Rev.* **7,** 25–39 (1977).
30. Layne, J. N. & Mckeon, W. H. *N.Y. Fish Game J.* **3,** 44–74 (1956).
31. Englund, J. *Viltrevy* **8,** 1–82 (1970).
32. Hewson, R. & Klob, H. H. *J. Zool.* **171,** 285–92 (1973).
33. Bogel, K., Moegle, H., Knorpp, F., Arata, A., Dietz, K. & Detthelm, P. *Bull. Wld Hlth Org.* **54,** 433–47 (1976).
34. Berger, J. *Lecture Notes in Biomaths* Vol. 11, 75–88 (Springer, Berlin, 1976).
35. Mollison, D. *J.R. statist. Soc.* **B39,** 283–326 (1977).
36. Preston, E. M. *J. Wildl. Mgmt* **37,** 501–12 (1973).
37. Sayers, B. Mcn., Mansourian, B. G., Phan Tan, T. & Bogal, K. *Med. Inform.* **1,** 11–34 (1977).
39. Frerichs, R. K. & Prawda, *J. Mgmt Sci.* **22,** 411–21 (1975).
40. Smart, C. W. & Giles, R. H. *Wildl. Dis.* **61,** 1–89 (1973).
41. Stellman, C. & Beranger, G. *Revue med. vet.* **125,** 45–64 (1974).
42. Lambinet, D., Boisvieux, J. F. Mallet, A., Artois, M. & Andral, L. *Rev. Epidem. Sante* **26,** 9–28 (1978).
43. Bacon, P. J. & MacDonald, D. W. *New Scient.* **87,** 640–645 (1980).
44. Anderson, R. M. & May, R. M. *Nature* **280,** 361–367 (1979).
45. May, R. M. & Anderson, R. M. *Nature* **280,** 455–461 (1979).
46. Scott, T. G. *Iowa State Coll. J. Sci.* **14,** 155–9 (1941).
47. Schoefield, R. D. *J. Wildl. Mgmt* **24,** 432–4 (1960).
48. Fairley, J. S. *Proc. R. Ir. Acad.* **69,** 103–37 (1970); *Ir. Nat. J.* **16,** 392–3 (1970).
49. Jensen, B. *Dan. Rev. Game Biol.* **8,** 1–20 (1973).
50. MacDonald, D. W. *Mammal. Rev.* **7,** 7–23 (1977).
51. Storm, G. L. *J. Wildl. Mgmt* **29,** 1–13 (1965).
52. Lloyd, H. G. *et al. Zentbl. VetMed.* **23,** 580–9 (1976).
53. Lloyd, H. G. & Englund, J. *J. Reprod. Fert. Suppl.* **19,** 119–130 (1973).
54. Ables, E. D. in *The Wild Canids—Their Systematics, Behavioural Ecology and Evolution* (ed. Fox M. V.) (Van Nostrand Reinhold, New York, 1975).
55. Wandeler, A. *Proc. Symp. Adv. Rabies Res.* (US Dept NIH Education and Welfare, Atlanta, 1976).
56. Englund, J. *Viltrevy* **3,** 377–485 (1965).
57. Verhulst, P. F. *Corresp. Math. Phys.* **10,** 113 (1938).
58. Bailey, N. T. J. *The Mathematical Theory of Infectious Diseases* 2nd edn (Macmillan, New York, 1975).
59. Black, J. G. & Lawson, K. F. *Can. J. comp. Med.* **34,** 309–311.

60. Anderson, R. M. & May, R. M. *Phil. Trans. R. Soc.* (in the press).
61. Dietz, K. in *Epidemiology* (eds Ludwig, D. & Cooke, K. L.) 104–121 (Society for Industrial and Applied Mathematics, Philadelphia, 1974).
62. Kermack, W. O. & Mckendrick, A. G. *Proc. R. Soc.* **A115,** 700–721 (1927).
63. Wld Hlth Org. *Report on a Conference Frankfurt-am-Main* WHO Regional Office for Europe, Copenhagen, 1978).
64. Muller, J. *Bull Off. int. épizoot.* **75,** 763–776 (1971).
65. Dean, D. J. *Publ. Hlth Rep., Wash.* **70,** 564–569 (1955).
66. Jennings, W. L., Schneider, N. J., Lewis, A. L. & Scatterday, J. E. *J. Wildl. Mgmt* **24,** 171–179 (1960).
67. Carey, A. B., Giles, R. H. & McLean, R. G. *Am. J. trop. Med. Hyg.* **27,** 573–80 (1978).
68. Schlegel, van H. L. *Dt. tierärztl. Wschr.* **83,** 311–350 (1976).
69. Kauker, E. & Zettl, K. *Berl. Münch. tierärztl. Wschr.* **16,** 301 (1969).
70. May, R. M. (ed.) *Theoretical Ecology: Principles and Applications* (Blackwell, Oxford, 1976).
71. Anderson, R. M. in *The Mathematical Theory of Biological Populations* (eds Hiorns, R. W. & Cooke, D. L.) (Academic, London, in the press).
72. Gross, E. C. *J. Mammal.* **21,** 294–306 (1940).
73. Elton, C. S. *Voles, Mice and Lemmings* (Clarendon, Oxford, 1942).
74. Taper, S. C. *Mammal. Rev.* **6,** 93–117 (1976).
75. Taper, S. C. *J. Anim. Ecol.* **48,** 603–617 (1979).
76. Davis, D. E. & Wood, J. E. *Publ. Hlth Rep.* **72,** 115–118 (1959).
77. Moegle, H. *et al. Zentbl. Vet. Reihe* **21,** 647 (1974).
78. Andral, L. & Toma, B. *Cah. Méd. vét.* **42,** 203–214 (1973).
79. Schaefer, M. B. *Inter-American Tropical Tuna Commission Bulletin,* 126–56 (1954).
80. May, R. M. *Nature* **269,** 471–477 (1977).
81. Beddington, J. R. & May, R. M. *Science* **197,** 463–465 (1977).
82. Bogel, K., Arata, A. A., Moegle, H. & Knorpp, F. *Zbl. vet. med.* **B21,** 401–12 (1974).
83. Turner, G. S. *Trans. R. Soc. trop. med. Hyg.* **70,** 175–8 (1976).
84. Blancou, J. *Recl. Med. Vet.* **155,** 733–41 (1979).
85. Black, J. G. & Lawson, K. F. *Can. Vet.* 3 **14,** 206–11 (1973).
86. Ramsden, R. D. & Johnston, P. H. *J. Wildl. Dis.* **11,** 318–24 (1975).
87. Baer, G. M., Broderson, J. R. & Yega, P. *Am. J. Epidemiol.* **101,** 160–4 (1975).
88. Winkler, W. G. & Baer, G. M. *Am. J. Epidemiol.* **103,** 408–15 (1976).
89. Manz, D. *Dt. Vet. Med. Ges., Bad Nauheim* **12** (4), 75 (1975).
90. Dubreuil, M., Andral, L., Aubert, M. F. & Blancou, J. *Ann. Rech. Vet.* **10,** 9–21 (1979).
91. Steck, F., Hafliger, U., Stocker, C. H. & Wandeler, A. I. *Experientia* **34,** 1662 (1978).
92. Winkler, W. G., McLean, R. G. & Cowart, J. C. *J. Wildl. Dis.* **11,** 382–388 (1975).
93. MacInnes, C. D. & Johnston, D. H. *Ontario Fish and Wildl. Rev.* **14,** 17–20 (1975).
94. Ministry of Agriculture, Fisheries and Food *Rabies Memorandum* (Ministry of Agriculture, Fisheries and Food and Department of Agriculture and Fisheries, Scotland, 1977).
95. *Rabies Bull. Europe* Vol. 3, Nos. 1–5 (WHO Collaborating Centre for Rabies Surveillance and Research, Tübingen, 1979).
96. Crick, J. & Brown, F. *Vet. Rec.* 28 Aug. 6–11 (1976).
97. Harris, S. *J. Zool. Lond.* **187,** 195–199 (1979).
98. Harris, S. thesis, University of London (1975).
99. Bromel, J. & Zettl, K. *Zeitschrift Jagdwissenshaft* **20,** 96–104 (1974).
100. Hauften J. Van *Trans. Int. Cong. Game Biologists* **9,** 539–543 (1970).
101. Zimen, E. (ed.) *The Red Fox. Symposium on Behaviour and Ecology* (Junk, The Hague, 1980).

The molecular structure and stability of the eye lens: X-ray analysis of γ-crystallin II

Tom Blundell, Peter Lindley, Linda Miller, David Moss, Christine Slingsby, Ian Tickle, Bill Turnell & Graeme Wistow

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK

The three-dimensional structure of the eye lens protein, bovine γ-crystallin II, has been determined at 2.6 Å resolution. The protein has a two domain β-structure, folded into four remarkably similar 'Greek key' motifs, and shows the highest internal symmetry of any protein studied by X-ray analysis. Although the symmetrical structure seems very stable, the arrangement of the sulphydryl groups would allow intramolecular cross-linking leading to possible destabilization, and intermolecular cross-linking leading to aggregation, both of which may be important in cataract formation.

ALTHOUGH eye lenses have evolved to suit the requirements of many ecological niches—from the very hard lens of the short-sighted rat to the very soft lens of birds capable of wide-ranging accommodation—all vertebrate lenses have a similar cellular structure comprising long fibre cells which pack hexagonally in spherical laminae. These grow throughout life, differentiating from the surface epithelial cells and overlaying older cells, so that the core of the adult lens is formed from the original fetal structure.

The transparency and refractive power of the normal lens depends on a smooth gradient of refractive index for visible light. This is achieved not only by the regular arrangement of the fibre cells, but also by a smoothly changing concentration gradient of lens-specific proteins, the crystallins, and other cell components.

The largest of the soluble crystallins, α-crystallin, is an oligomeric globular protein of molecular weight (MW) $\sim$800,000. There are two gene-product subunits, αA_2 and αB_2, both of MW[1,2,3] 20,000. A proportion of these undergo post-translational modification[4], which contributes to the considerable heterogeneity and polymerization of native α-crystallin[5,6]. β-Crystallins are also a heterogeneous class of proteins which exist as oligomers of several different sizes[7,8]; the subunits vary in MW between 20,000 and 32,000 with the predominant species, βBp, of MW 26,000 (ref. 9). γ-Crystallins, however, are monomers of 20,000 MW[10], with highly homologous sequences[11,12]. They show no similarity to α-crystallin, but a striking homology to the sequence of βBp-crystallin has been recently observed[13]. γ-Crystallin II, the major species in the bovine lens, has six free sulphydryl residues[11].

Little is known of the conformation and intermolecular interactions of the crystallins which give rise to the supramolecular organization within the fibre cells. However, it is established that the increased light scattering of senile nuclear cataracts is associated with protein aggregation[14–16], together with chemical cross-linking, including disulphide bridges[17–20] and denaturation[21].

Here we report the three-dimensional structure of bovine γ-crystallin II at 2.6 Å resolution—the first X-ray analysis of a lens protein. The protein has a two-domain β-structure, folded

Table 1 Phasing statistics for the five heavy-atom derivatives given in resolution ranges

Reagent	Uranyl acetate	Thiomersal		Potassium auricyanide		Ethyl mercury chloride	Sodium hexachloroiridate (IV)	
Resolution	16→4.5 Å	16→3.4 Å	3.4→2.6 Å	16→3.4 Å	3.4→2.6 Å	16→3.4 Å	16→3.3 Å	3.3→2.8 Å
$\bar{F}_{PH}$	323.7	321.1	209.4	318.8	208.5	296.8	313.9	220.2
$\bar{F}_H$	35.7	64.9	44.3	51.7	41.9	100.1	54.4	40.2
E_H	40.9	37.2	35.0	39.5	37.4	66.6	60.1	43.0

$\bar{F}_{PH}$, mean value for the derivative protein structure factor amplitude. $\bar{F}_H$, mean value of the contribution of the heavy atom to the structure factor of the derivative. E_H, r.m.s. lack of closure[27] for the heavy-atom derivative.

0028-0836/81/090771—07$01.00

		←		a		→							←			b		→			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
γ II Motif I	GLY	LYS	ILE	THR	PHE	TYR	GLU	ASP	ARG	GLY	PHE	GLN	GLY	HIS	CYS	TYR	GLN	CYS	SER	SER	ASN
		←		β		→	R	← β	→	← L	→	β	G	←	β		→				
	40	41	42A	43A	44A	45A	47	48	49	50	51	51A	52	53	55	55A	55B	55C	55D	56	57
Motif II	HIS	SER	TRP	PHE	VAL	TYR	GLN	ARG	PRO	ASP	TYR	ARG	GLY	HIS	GLN	TYR	MET	LEU	GLN	ARG	GLY
		←		β		→	R	← β	→	← L	→	β	G	←		β		→			
		80	81	82	83																
Connecting peptide		HIS	THR	GLY	THR																
	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	
Motif III	PHE	ARG	MET	ARG	ILE	TYR	GLU	ARG	ASP	ASP	PHE	ARG	GLY	GLN	MET	SER	GLU	ILE	THR	ASP	
	←			β		→	R	← β	→	← L	→	β	G	←		β		→			
	123	124	125	125A	125B	125C	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140
Motif IV	GLY	SER	TRP	VAL	ILE	TYR	GLU	MET	PRO	SER	TYR	ARG	GLY	ARG	GLN	TYR	LEU	LEU	ARG	PRO	GLY
	←			β		→	R	← β	→	← L	→	β	G	←			β			→	

← c	→																←			d				→			
22	23	24	25	26A	27	28				29	30					31	32	33	34	35	36		37	38	39		
ASN	CYS	LEU	GLN	GLN	PRO	TYR				PHE	SER					ARG	CYS	ASN	SER	ILE	ARG		VAL	ASP	VAL		
← β	→	←		R		→										← β	→			←		β		→			
58	59	60	61	62	62A	63				64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79		
ASN	TYR	PRO	GLN	TYR	GLY	GLN				TRP	MET	GLY	PHE	ASP	ASP	SER	ILE	ARG	SER	CYS	ARG	LEU	ILE	PRO	GLN		
← β	→	←		R		→						← R			→	← β	→			←	β		→				
104A	105	106A	107	108	109	110	111	112	113	114	115					116	117	118	119A	119B	119C		120	121	122		
ASP	CYS	PRO	SER	LEU	GLN	ASP	ARG	PHE	HIS	LEU	THR					GLU	VAL	ASN	SER	VAL	ARG		VAL	LEU	GLU		
← β	→ β	←			R				→							← β	→			←	β		→				
141	142	143	144	145	146	147				148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165
GLU	TYR	ARG	ARG	TYR	LEU	ASP				TRP	GLY	ALA	MET	ASN	ALA	LYS	VAL	GLY	SER	LEU	ARG	ARG	VAL	MET	ASP	PHE	TYR
← β	→	←		R		→						←	R		→	← β	→			←	β		→				

Fig. 1 The sequence of γ-crystallin II. The numbering is that of Croft[11] and subscripts indicate insertions or replacements necessitated by the electron density map. These generally improve the homology with the sequence of βBp-crystallin reported by De Jong *et al.*[13]. The sequence is divided into four major, approximately equal parts which correspond to the four Greek key motifs, each with four strands: a, b, c and d. Topologically equivalent residues are in columns and the secondary structure indicated by the torsion angles is given: β, β-sheet; R, right-handed helix generally distorted from α-helix; L, left-handed helix, and G, a conformation found usually with glycine, with $\phi \sim 110°$ and $\psi \sim -110°$. The connecting peptide links the two domains.

into four remarkably similar Greek key motifs[22]; it shows the highest internal symmetry of any protein studied by X-ray analysis. Each domain comprises a sandwich of two four-stranded sheets, each sheet being made from three strands of one 'Greek key' motif and one strand from another. The sequence shows a functional conservation of a few important residues in each of these motifs which had not been recognised earlier. This repeating structure, which may have evolved through gene multiplication, could be of general significance to other lens proteins in view of the sequence homology with βBp-crystallin[13]. Although the symmetrical structure has the appearance of great stability, the sulphydryl groups are arranged in ways which could allow both intramolecular and intermolecular cross-linking, and reaction with certain sulphydryls could be expected to destabilize the conformation and lead to denaturation. The arrangement of charged and non-polar groups on the molecular surface has implications for the stability of the molecule, the nature of the hydration around lens proteins and the intermolecular contacts which may be important to the supramolecular organization of crystallins in the normal lens.

X-ray analysis

The γ-crystallins were isolated from calf lens using gel filtration on G-75 Sephadex to separate the α- and β-crystallins[23], ion-exchange chromatography on DEAE-Sephadex to separate the βs-crystallin (MW 28,000)[24], and ion-exchange chromatography on sulphopropyl-Sephadex and DEAE-cellulose to separate the γ-crystallins I, II, III and IV (refs 25, 26). Large crystals of γ-crystallin II were grown from 0.05M phosphate buffer, 1mM dithiothreitol, *p*H 7, at 0 °C (ref. 26). The crystals were stabilized by cross-linking overnight in 3% glutaraldehyde in 0.05M sodium phosphate buffer, *p*H 7. Four heavy-atom derivatives were prepared using uranyl acetate, thiomersal, potassium auricyanide and ethyl mercury chloride as described elsewhere[27] and a further derivative using sodium hexachloroiridate (IV). Data were collected on a Hilger-Watts four-circle diffractometer with Ni-filtered CuK$_\alpha$ radiation to the resolutions indicated in Table 1. Typically, ω scans with 1° scan width and 50-s counting time including background were analysed using a moving-window technique[28] to obtain the net integrated intensity. Two equivalents (hk*l* and kh*l*) for the native and four equivalents (hk*l*, kh*l*, hk$\bar{l}$, kh$\bar{l}$) for the derivatives were measured so that two estimates of the anomalous difference of each reflection for the derivatives were obtained. Empirical absorption corrections were applied using a modification of the method of North *et al.*[29] and radiation decay was monitored by standard reflections measured every hour. The heavy-atom positions were determined using F_{HLE} Patterson syntheses and refinement (see ref. 30 for review) and the relative hands and origins of the individual derivatives, and the absolute hand of the model, were established using cross-phase difference Fourier syntheses incorporating the anomalous differences[27]. Phases were calculated using the method of Blow and Crick[31] and the phasing statistics are summarized in Table 1. The mean figure of merit was 0.89 at 5.5 Å resolution but fell to 0.70 in the range 3.0–2.6Å resolution.

A 'mini' electron density map at 3.4 Å resolution was calculated at a scale of 1 cm = 3Å and the path of the polypeptide was followed. The 2.6 Å electron density map was initially interpreted using an optical comparitor[32] (scale 1.25 cm = 1 Å) and the α- and β-carbon, carbonyl oxygen and some side-chain positions of this model were recorded. These were then displayed on an Evans and Sutherland Picture System II using FRODO[33] modified and extended by A. Jones and I. J. Tickle. The backbone atoms were generated using the input coordinates as guide points, with the model-building routine of Hermans[34], in the FRODO program and the coordinates improved interactively using computer graphics. The addition of side chains presented problems as some modifications to the sequence published by Croft[11] were required mainly in the hydrophobic regions. Our sequence (Fig. 1) is an attempt to fit the observed electron density but at the same time to minimize the number of implied errors in tryptic and cyanogen bromide peptide sequences which were used to derive the primary structure. The new sequence has higher homology to the sequence of βBp (ref. 13) and also higher internal homology (*vide infra*) than that published previously[11]. The numbering scheme uses alphabetical postscripts where the residues differ from the Croft sequence.

The topology of the folding units was compared using EZIFIT of A. D. McLachlan[35]. Thirty-six topologically equivalent residues were compared, including all residues shown in Fig. 1, and equivalent to residues 1–37 with the exception of residue 21. This program optimizes the fit without assuming a strict 2-fold rotational symmetry. A program written independently by W.G.T. was used to find the agreement to a rotation of 180°. The motifs were also compared using interactive computer graphics to optimize their overlap, using a program FITZ written by Dr G. Taylor of our laboratory.

Symmetry and stability

The low-resolution electron density map[27] showed that the γ-crystallin molecule is ellipsoidal and organized into two well defined globular domains. It can now be seen that the N-terminal half of the sequence comprizes one domain and the C-terminal half the second; the overall topology of the two domains is strikingly similar. Each domain is organized into two folding units (see Figs 2–4) that adopt a structure resembling a 'Greek key' motif[22] shown schematically in Fig. 3*b*. We have divided the sequence into four sections of about 40 residues, labelled I–IV in Fig. 1, so that I and II correspond to the Greek key motifs in the N-terminal domain and III and IV correspond to the motifs of the C-terminal domain. There is a short connecting peptide between the two domains, which are related by an approximate diad axis so that 80 α-carbon atoms of the N-terminal domain (38 of motif I and 42 of motif II) are topologically equivalent to those of the C-terminal domain with root mean square (r.m.s.) distance at best fit of 1.42 Å as calculated using EZIFIT[35] (rotation angle, 178.1°; screw shift, 0.14 Å).

We compared 36 common, topologically equivalent α-carbon atoms of each motif. For motifs I and II related by an approximate diad within the N-terminal domain, the r.m.s. distance was 2.40 Å (rotation angle, 179.3°; screw shift, 1.32 Å). Motifs III and IV in the C-terminal domain similarly gave a r.m.s. distance of 2.76 Å (rotation angle, 179.8°; screw shift, 1.12 Å). However, the most striking relationship is between motifs II and IV, where 42 consecutive α-carbon atoms in the two motifs are related with a r.m.s. distance of 1.05 Å, and between motifs I and III where 36 topologically equivalent residues are related with a r.m.s. distance of 1.74 Å.

The diad relating atoms of the two domains bisects the angle of ~40° between the diads relating folding units, all three diads being approximately coplanar. These observations demonstrate that folding units I and III are similar, as are II and IV. There is a less close correspondence between the folding units within one domain—this is illustrated in Fig. 2 which compares the α-carbons of the folding units.

Given the similarities between the Greek key motifs it is convenient to first describe motif I and then to relate the others to this. The sequence in Fig. 1 has been arranged so that topologically equivalent residues in each motif occur in one column, and an indication of the conformation of residues is included so that secondary structures can be inferred.

Each 'Greek key' comprises four antiparallel strands. In motif I (Fig. 3*a*), residues 2–6 (strand a) form an extended chain in β-sheet conformation that runs antiparallel and is hydrogen-bonded to residues 14–18 (strand b). The turn between residues 7 and 13 is made possible by Glu 7 assuming a right-handed helical conformation and Gly 13 having the ϕ, ψ values of 126° and −116°, respectively, which are normally disallowed for residues with side chains. Residues 10 and 11 assume the conformation of a left-handed α-helix so that the side chain of

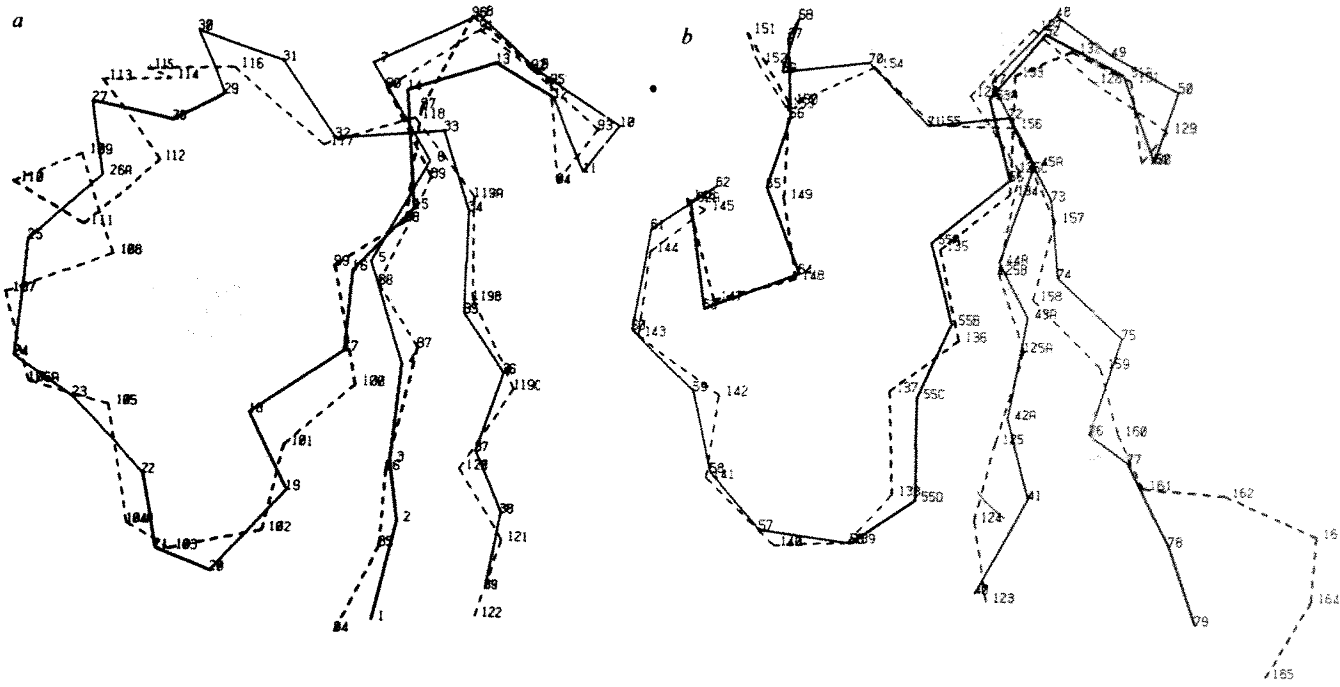


Fig. 2 The path of the polypeptide chain is illustrated by the positions of the α-carbon atoms for each of the folding motifs defined in Fig. 1. *a* Shows the best fit of motifs I and III obtained visually using interactive computer graphics (program FITZ of Dr G. Taylor). *b*, A similar best fit from an equivalent direction for motifs II and IV, which show the closest resemblance in sequences between the topologically equivalent residues.

Phe 11 is packed on to the antiparallel sheet against Tyr 6. The structure of residues 1–18 may therefore be described as a two-stranded antiparallel β-pleated sheet which has a turn at residues 10 and 11 and which is folded at residues 7 and 13.

Residues 19–21 form a further turn so that residues 22 and 23 lie antiparallel and adjacent to residues of strand b. After residues 22 and 23, which have a β-sheet conformation, the chain assumes a rather irregular, right-handed helical structure between residues 24 and 28. The following residues form a gentle bend so that residues 32–38 form strand d, which is hydrogen-bonded as part of the antiparallel β-sheet including strands a and b.

Figures 2–4 indicate that the conformations of motifs II, III and IV may be derived from that of motif I by insertions and small changes in conformation. The major differences between motifs I and III are the deletion of one residue in the turn between strands b and c and the insertion of three residues as a continuation of the right-handed helical region at the end of strand c. Motifs II and IV are slightly larger than I, with four insertions in the loop between the helical region at the end of strand c and strand d. Figure 4 shows that the hydrogen bonding in the β-sheets is similar.

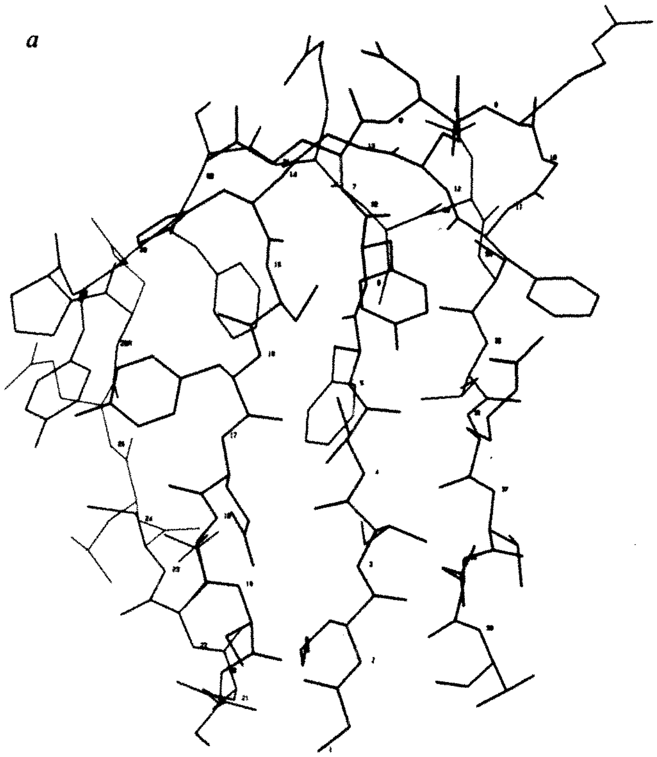
The Greek key motifs pack together so that each domain comprises two right-handed, twisted, antiparallel β-sheets sandwiched together (Fig. 3*c*). The fourth strand of each sheet is comprised of strand c of the other motif in the same domain as indicated in Fig. 4. As strand c is of a less regular β-sheet structure than the other strands, it makes few well defined hydrogen bonds. The angle between the planes of the β-sheets is ~35° giving the domains the wedge shape seen in Fig. 3*c*, and the angle between individual strands in opposing sheets projected on to the mean plane is ~20°. The volume between the β-sheets is occupied by hydrophobic side chains; these are more predominantly aromatic near the top of the wedge. The top of the wedge-shaped domain is enclosed by the two antiparallel strands c and d of the two Greek key motifs. In motif I the following residues contribute to the core: 3, 5, 16, 18, 23, 28, 32, 35, 37; the topologically equivalent residues in Fig. 1 (columns beneath the residues listed) also contribute to the cores of the domains, making them compact and largely hydrophobic. Residues in strands b and c have their hydrophilic functional groups folded in such a way as to minimize the contact of the core with the solvent.

The two domains are packed together (Fig. 5) so that the two sheets involving strands a, b and d of motifs II and IV are sandwiched with an angle of 20° between the individual strands. The contacts are again largely hydrophobic and include Phe 43A and Met 55B of the N-terminal domain, and Val 125A and Leu 136 of the C-terminal lobe. The connecting peptide is accommodated so that it follows the surface of the molecule.

The structure of γ-crystallin shows the highest degree of intrachain symmetry of any protein whose three-dimensional structure has been elucidated. The eye lens proteins must exist throughout the life of the animal without denaturation, due to the fact that to avoid sharp changes in refractive index which would result from irregular structures, the fibre cells become denucleated and have few cellular organelles such as mitochondria and ribosomes. As there is little or no capacity for regeneration, the proteins must be relatively stable, and we suggest that the high degree of symmetry of these proteins may contribute towards this stability.

Gene multiplication in evolution?

The similarity in the three-dimensional structure of the motifs poses the question as to whether the structure may have evolved divergently in two stages: first, gene duplication and fusion of a protein of ~40 amino acid residues to a protein approximating the present γ-crystallin domain, followed by a further gene duplication to give the two-domain structure. This model is consistent with the observation that the two domains resemble each other more closely than do the two motifs within each domain. This is also evident from the sequence homologies



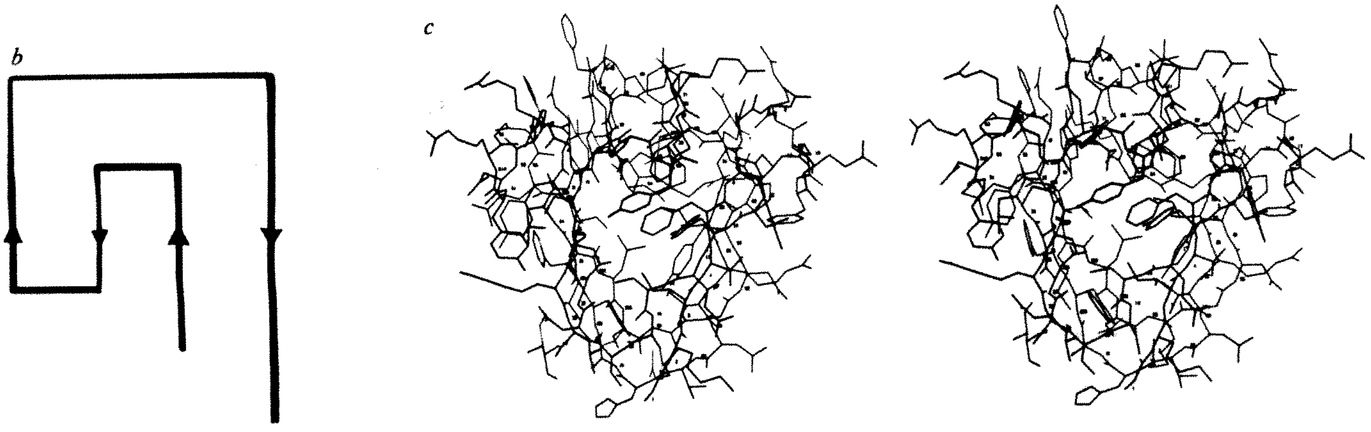


Fig. 3 *a*, The arrangement of the amino acid residues in a Greek key folding motif viewed at about right angles to the plane formed by strands a, b and d, which are defined in Fig. 1. *b*, The Greek key motif shown schematically. *c*, Stereo diagram of the N-terminal domain showing the wedge-shaped packing of the two motifs.

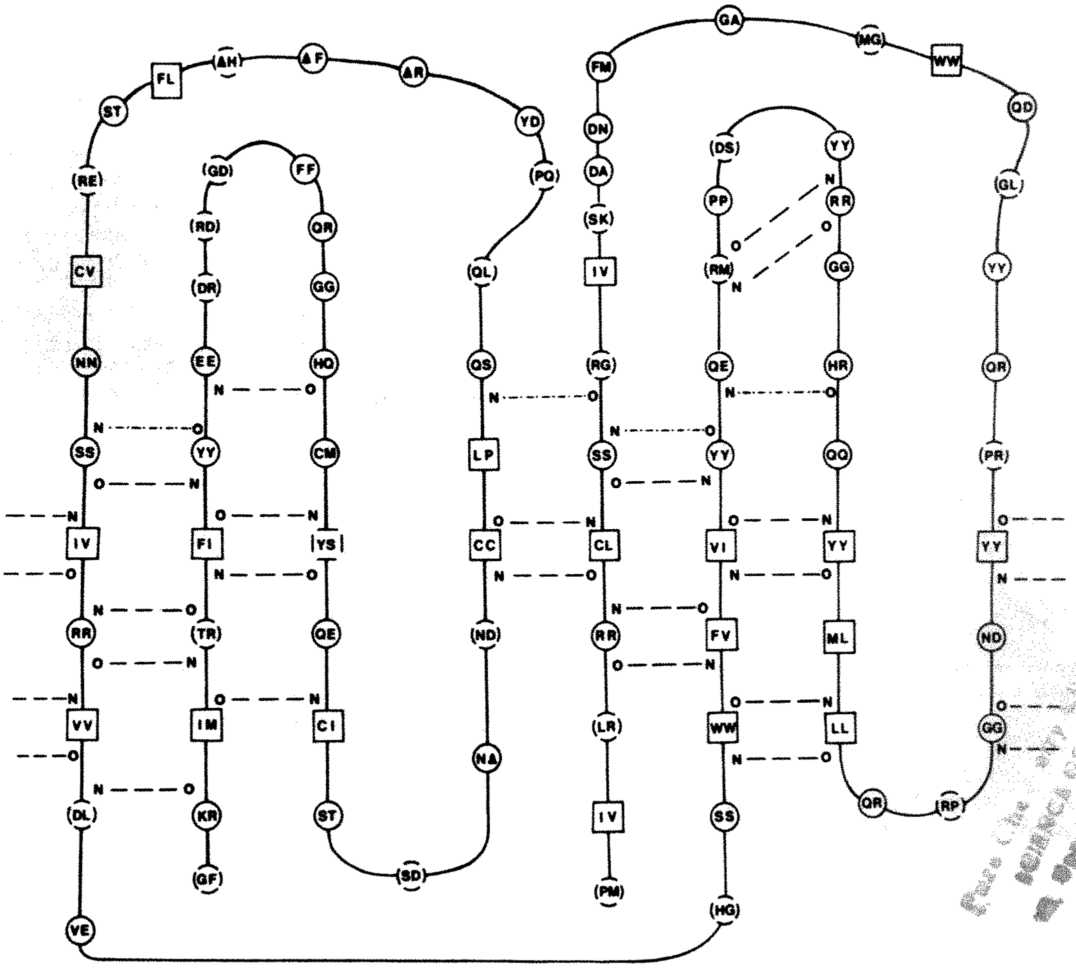
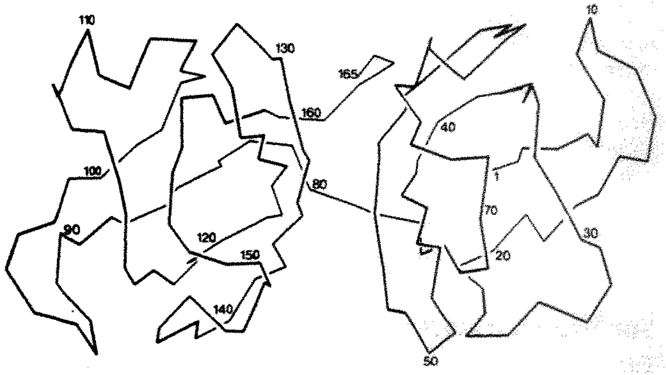


Fig. 4 The arrangement of the amino acids in a γ-crystallin domain shown schematically. Each box indicates the nature of the two topologically equivalent amino acids of the two domains. Deletions in one domain are given by ▲. Square boxes indicate residues which are directed either between the sheets within the domains or between the sheets between the domains; some of these residues in strands b and c (see Fig. 1 for definition) are on the edge of the sheets and are therefore partially accessible to solvent. The extent of the sheets is indicated by the length of strands shown in Fig. 1 and also by the hydrogen-bonding network. Those hydrogen bonds common between the two domains are indicated by ----- whereas those occurring in one domain only are indicated by -·-·-.

illustrated in Figs 1 and 4, which must, however, be read with caution as they involve inferences from the electron density map. The number of residues conserved at topologically equivalent positions between motifs is much greater for motifs II and IV than for the two motifs within either domain. These observations are consistent with the model for two sequential tandem duplications.

Only four residues are invariant at topologically equivalent positions between all folding units. These include Tyr 6 which has a structural role in packing against the conservatively varied Phe 11 (Tyr) of the folded β-turn between strands a and b (Fig. 6) and Gly 13 which has angles disallowed for other residues and allows the fold to occur. Glu 7, which is necessary for the folding

a



b

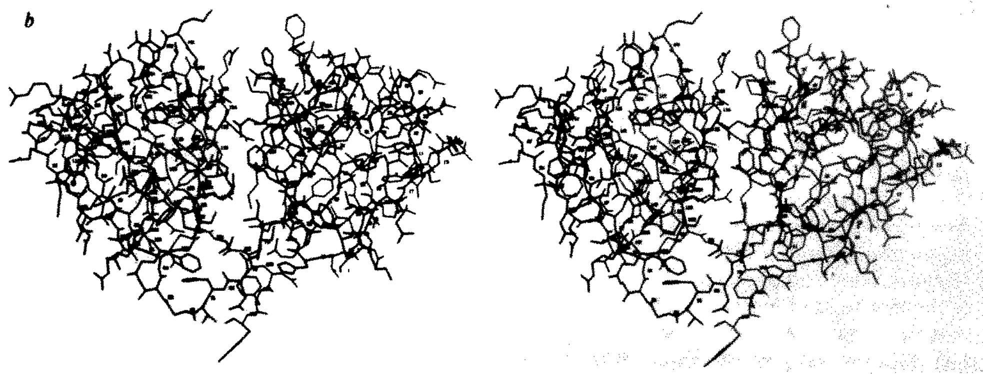


Fig. 5 The arrangement of the two domains in the molecule of γ-crystallin II. *a*, The C^{α} backbone viewed almost parallel to the pseudo 2-fold axis between the domains which are clearly shown. *b*, A stereo diagram of the complete molecule viewed approximately perpendicular to the plane, containing the three pseudo 2-fold axes.

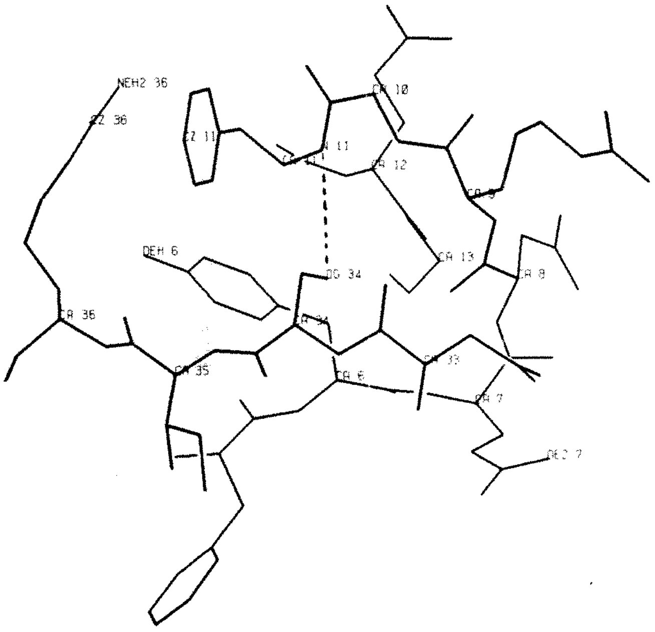


Fig. 6 The portion of motif I in the N-terminal domain containing residues invariant at topologically equivalent positions in the four motifs.

over of the β-structure, is very conservatively varied to Gln (in motif III alone). The third invariant residue, Ser 34, is also involved in this structure (see Fig. 6). The γ-O of this serine in each folding unit is relatively shielded from solvent and hydrogen bonds to the peptide nitrogen of the residue topologically equivalent to Phe 11 on the turn. This most unusual structural motif is further stabilized by the fourth invariant residue Arg 36, also shown in Fig. 6; this arginine seems to be well ordered and to have its aliphatic side chain against the aromatic ring of Phe 11 with its guanidinium group exposed to the solvent. The fact that this arrangement is strongly conserved both in three dimensions and in sequence underlines the importance of these amino acids, but their structural role would be consistent with convergent evolution as well as gene multiplication and divergent evolution.

One of the surprising aspects of the structure is that very few residues of the hydrophobic core are invariant. It seems that there are several ways in which the same volume occupation of side chains of the core can be achieved while still allowing a very close three-dimensional homology. Evolution–whether convergent or divergent–seems to have exploited these possibilities and it is this variation, combined with the insertions and deletions in the bends, which makes recognition of structural homology very difficult from sequence alone. If this is divergent evolution, here is further evidence that three-dimensional structure is more strongly conserved than sequence.

Surface characteristics

The surface of the γ-crystallin II molecule is characterized by a number of extremely well defined charged and polar side chains, many of which seem to be organized in pairs or chains of ionic and/or polar interactions. For example, Glu 7 and Asp 8 at the beginning of the fold between strands a and b of motif I have their acidic side chains lying on either side of the basic guanidinium moeity of Arg 31 at the beginning of strand d; Glu 116 is bracketed by Asn 118 and Asn 152; Glu 122, Arg 85, Glu 100 and Arg 87 form a chain of ionic interactions.

The most extensive network of charged groups is centred on Asp 92 and includes a series of six alternately positively and negatively charged side chains in one molecule (Fig. 7). However, one end of the chain, Arg 144, is close to Asp 103 of a symmetry-related molecule at a crystal lattice contact point so that the ionic interactions connect neighbouring molecules. In fact, Glu 90 and Arg 91, adjacent to Asp 92 but probably not linked ionically, form an ion pair dimer with their symmetry-related equivalents Arg 91 and Glu 90 across a 2-fold axis to a third molecule in the lattice.

It seems likely that at the intramolecular level, these interactions are partially responsible for the high stability that seems to be functional in this structure. It has been pointed out[36] that enzymes from thermophilic bacteria have abundant surface ion pairs which contribute to thermal stability. An 'exoskeleton' of organized charged groups could therefore be quite important in holding together the strands that form each domain of γ-crystallin II.

On the level of intermolecular contacts, not only hydrophilic residues are involved. The few relatively exposed hydrophobic side chains, such as Phe 11 and Tyr 51 are often found at lattice contact points (Phe 11 at a 2-fold near its symmetry-related equivalent). Similar interactions may occur in the normal lens and contribute to its supramolecular organization and control of protein concentrations. Differences in the balance of these polar and non-polar surface residues may be responsible for the different properties of the various γ-crystallins, for example, their relative abilities to cryoprecipitate. A comparative study of γ-crystallins is now being done in our laboratory to investigate these effects (parallel work is being done by Dr Y. Chirgadze[37,38]).

The nature of the molecular surface may also be important when considering the structure of the water in the lens. Eye lenses generally have a very low water content, particularly in the most central regions where, in mammals, γ-crystallins are most concentrated. The percentage of water correlates with the refractive power of the lens and its degree of softness and therefore its ability to accommodate.The charge interactions that fix the positions of polar side chains, making them so well defined in the electron density map, would lead to partial neutralization of charge, but by being fixed, the surface ions could still have a major effect in ordering local water. One of the roles of the crystallins may be to form hydration spheres and regulate the water content of the lens.

Sulphydryl arrangement and accessibility

It is surprising that a protein which must retain its conformation and seems to have a largely structural role should have evolved with six free sulphydryl groups . An intracellular protein of this size would be expected to have about two cysteine residues[39]. It was therefore interesting to examine the accessibility, reactivity and arrangement of these groups.

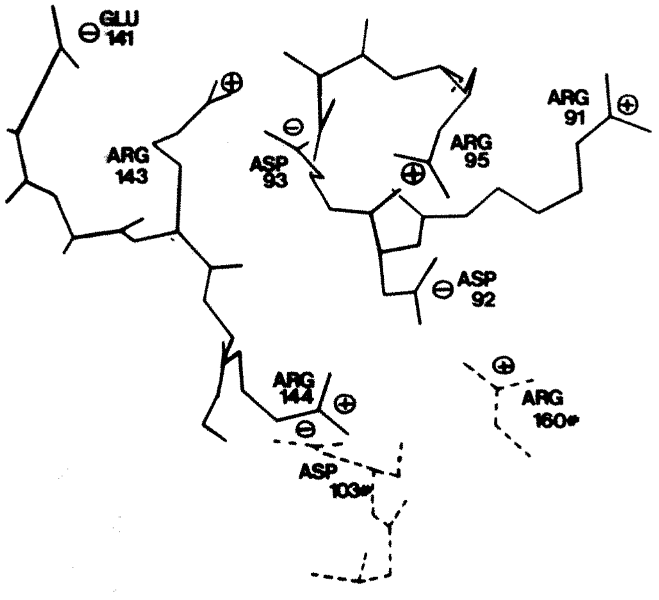


Fig. 7 An extensive network of charged groups on the surface of the molecule, centred on Asp 92. This and similar interactions may be partially responsible for the stability of the γ-crystallin molecule.

The three-dimensional structure demonstrates that not only do five of these groups lie in the same domain but three are also clustered closely together. Figure 4 indicates that Cys 23 and Cys 74 lie together on the same side of adjacent strands of a β-sheet in a position to form a disulphide bridge. Cys 18 also lies close by on the other sheet in the same domain, and this might also form an intramolecular disulphide bridge with Cys 23 in non-reducing conditions. Such intramolecular disulphide bridges would not radically affect the overall conformation of the molecule. Cys 15 and Cys 105, like Cys 23 and Cys 18, are on either strand b or c of the Greek key motif, and are therefore accessible to solvent. In the appropriate oxidizing conditions, these cysteines could form intermolecular disulphide bonds. In species such as the rat, where the core lens proteins undergo almost complete oxidation of protein sulphydryl[40,41], a combination of inter- and intramolecular disulphide cross-links must occur, involving such surface and clustered cysteines.

The two remaining cysteine residues, 32 and 74, are relatively buried in the structure. We observed early on in our study that extensive reaction in the crystals with small, reactive groups such as ionic mercury salts led to a rapid loss of isomorphism, and that reaction in solution led to precipitation. This did not occur if the mercury was coordinated with softer anions, attached to large groups (as in thiomersal) or part of a small hydrophobic cation like $EtHg^+$. The buried cysteine residues are not accessible to the large thiomersal group, whereas the small $EtHg^+$ can bind to both, causing only small changes in conformation, as the ethyl group does not greatly disturb the hydrophobic core of the molecule. It is probable that Hg^{2+} ions bind to Cys 32 and 74 but then, by production of $CysHg^+$, cause destabilization and possibly denaturation of the molecule. Such an irreversible covalent change leading to denaturation could be the first stage in a series of catastrophic and cascading events leading to the extensive polymerization and denaturation observed in senile nuclear cataracts. Whether these observations are relevant to human senile cataract must, however, await the chemical characterization of human γ-crystallin (Augusteyn *et al.*, unpublished results).

We have thus shown that although γ-crystallin II has a remarkably symmetrical structure which may confer stability on this protein in the normal healthy lens, the unusual arrangement of sulphydryls could contribute to the process of cataract if the reducing environment of the lens, including glutathione levels, were disturbed. Furthermore, the existence of a highly organized arrangement of charges on the molecular surface may be important in the organization of the crystallins and water in the normal transparent lens.

We thank Professor H. Bloemendal and Drs W. W. de Jong and L. Croft for sequence information on the βBp- and γ-crystallins respectively, and Dr J. Harding for useful and stimulating discussions. This research was supported by MRC grant no. 76/862.

Received 12 September; accepted 27 November 1980.

1. Van der Ouderaa, F. J., de Jong, W. W. & Bloemendal, H. *Eur. J. Biochem.* **39,** 207–222 (1973).
2. Van der Ouderaa, F. J., de Jong, W. W., Hilderink, A. & Bloemendal, H. *Eur. J. Biochem.* **49,** 157–168 (1974).
3. Spector, A. & Katz, E. *J. biol. Chem.* **240,** 1979–1985 (1965).
4. Van Kleef, F. S. M., de Jong, W. W. & Hoenders, H. J. *Nature* **258,** 264–266 (1975).
5. Siezen, R. J., Bindels, J. G. & Hoenders, H. J. *Eur. J. Biochem.* **91,** 387–396 (1978).
6. Stauffer, J., Rothschild, C., Wandel, T. & Spector, A. *Invest. Ophthalmol. Vis. Sci.* **13,** 135–146 (1974).
7. Herbrink, P., Van Westreenen, H. & Bloemendal, H. *Expl Eye Res.* **20,** 541–548 (1975).
8. Asselbergs, F. A. M., Koopmans, M., Van Venrooij, W. J. & Bloemendal, H. *Expl Eye Res.* **28,** 223–228 (1979).
9. Bloemendal, H. & Herbrink, P. *Ophthal. Res.* **6,** 81–92 (1974).
10. Björk, I. *Expl Eye Res.* **9,** 152–157 (1970).
11. Croft, L. R. *Biochem. J.* **128,** 961–970 (1972).
12. Slingsby, C. & Croft, L. R. *Expl Eye Res.* **26,** 291–304 (1978).
13. Driessen, H. P. C., Herbrink, P., Bloemendal, H. & de Jong, W. W. *Expl Eye Res.* **31,** 243–246 (1980).
14. Philipson, B. T. & Fagerholm, P. P. *Ciba Fdn Symp.* **19,** 45–63 (1973).
15. Jedziniak, J. A., Nicoli, D. F., Baram, H. & Benedek, G. B. *Invest. Ophthalmol. Vis. Sci.* **17,** 51–57 (1978).
16. Spector, A. *et al. Science* **204,** 1323–1326 (1979).
17. Anderson, E. I. & Spector, A. *Expl Eye Res.* **26,** 407–417 (1978).
18. Takemoto, L. J. & Azari, P. *Expl Eye Res.* **23,** 1–7 (1976).
19. Truscott, R. J. W. & Augusteyn, R. C. *Expl Eye Res.* **25,** 139–148 (1977).
20. Spector, A. & Roy, D. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3244–3246 (1978).
21. Harding, J. J. & Dilley, K. J. *Expl Eye Res.* **22,** 1–73 (1976).
22. Richardson, J. S. *Nature* **268,** 495–500 (1977).
23. Björk, I. *Expl Eye Res.* **1,** 145–154 (1961).
24. Van Dam, A. F. *Expl Eye Res.* **5,** 255–266 (1966).
25. Björk, I. *Expl Eye Res.* **3,** 254–261 (1964).
26. Carlisle, C. H., Lindley, P. F., Moss, D. S. & Slingsby, C. *J. molec. Biol.* **110,** 417–419 (1977).
27. Blundell, T. L. *et al. Acta crystallogr.* **34,** 3653–3657 (1978).
28. Tickle, I. J. *Acta crystallogr.* **31,** 329–331 (1975).
29. North, A. C. T., Phillips, D. C. & Mathews, F. S. *Acta crystallogr.* **24,** 351–359 (1968).
30. Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, London, 1976).
31. Blow, D. M. & Crick, F. C. *Acta crystallogr.* **12,** 794–802 (1959).
32. Richards, F. M. *J. molec. Biol.* **37,** 225–230 (1968).
33. Jones, T. A. *J. appl. Crystallogr.* **11,** 268–272 (1978).
34. Hermans, J. & McQueen, J. E. *Acta crystallogr.* **30,** 730–739 (1974).
35. McLachlan, A. D. *J. molec. Biol.* **128,** 49–80 (1979).
36. Walker, J. E., Wonacott, A. J. & Harris, J. I. *Eur. J. Biochem.* **108,** 581–586 (1980).
37. Chirgadze, Iu. N., Nikonov, S. V., Garber, M. B. & Reshetnikova, L. S. *J. molec. Biol.* **110,** 619–624 (1977).
38. Chirgadze, Iu. N., Sergeev, Iu. V., Fomenkova, N. P., Nikonov, S. V. & Lunin, V. Iu. *Proc. Acad. Sci. USSR* **250,** 762–765 (1980).
39. Fahey, R. C., Hunt, J. S. & Windham, G. C. *J. molec. Evol.* **10,** 155–160 (1977).
40. East, E. J., Chang, R. C. C., Yu, N. T. & Kuck, J. F. R. Jr., *J. biol. Chem.* **253,** 1436–1441 (1978).
41. Askren, C. C., Yu, N. T. & Kuck, J. F. R. *Expl Eye Res.* **29,** 647–654 (1979).

LETTERS

Constraints on heavy neutrinos

Doug Toussaint & Frank Wilczek

Institute for Theoretical Physics, University of California,
Santa Barbara, California 93106, USA

If a heavy neutrino of mass 1.1 MeV < m < 14 MeV participates in the weak charged current coupling for the electron, then it can be produced in the Sun and decay in flight into an electron, a positron and a light neutrino. It is shown here that existing experimental bounds on low-energy interplanetary positrons severely constrain the heavy neutrino–electron coupling. Accelerator experiments severely constrain the coupling of still heavier neutrinos.

Several recent developments in high-energy physics and in astrophysics have reawakened interest in the question of neutrino masses. There have been experiments on the end-point energy in tritium decay[1] and on neutrino oscillations[2] which purport to show positive results. The tritium experiment is easier to quantify, and if correct shows that at least one of the neutrinos emitted in tritium decay with large amplitude has a mass of a few tens of electron volts. Also, some unified theories of strong, electromagnetic and weak interactions suggest that neutrinos are massive[3]. In these theories leptons are treated in parallel with quarks, except that the right-handed components of the neutrinos are $SU(3) \times SU(2) \times U(1)$ singlets (having neither strong, electromagnetic nor weak interactions) and so can acquire a large Majorana mass. In these theories we often find mass formulae of the form

$$m_\nu \propto m_q^2/m_N \qquad (1)$$

where m_ν is the neutrino mass, m_q is the mass of the corresponding charge 2/3 quark (for example, u quark for the lightest neutrino, c quark for the next lightest) and m_N is the Majorana mass. The quark–lepton analogy might also suggest small mixing angles among different neutrinos, analogous to the Cabibbo angle for quarks. Finally, observational evidence for non-luminous matter has become quite strong[4], reviving interest in suggestions that massive neutrinos could contribute significantly to the overall mass density in the Universe[5] and in

0028–0836/81/090777—02$01.00

galaxies and their haloes[5,6]. In these regards it is neutrinos with mass of a few tens of electron volts, and lifetimes longer than the age of the Universe, that are most significant. Heavier stable or quasi-stable neutrinos are ruled out by these astrophysical considerations. (However, if the neutrino mass gets above several GeV, these limits no longer apply, as noted by Lee and Weinberg[7].)

These considerations suggest that the lightest neutrino, coupling mainly to the electron, may have a mass of a few tens of electron volts. According to equation (1) we would expect the other neutrinos to be much heavier, and presumably unstable on cosmological time scales due to some coupling with the electron charged current through mixing. We will argue that existing observational evidence on the abundance of low-energy interplanetary positrons severely constrains the allowable mixing angles for neutrinos with mass $1.1\,\mathrm{MeV} < m < 14\,\mathrm{MeV}$. (Astrophysical constraints on very long-lived neutrinos which decay radiatively have been given by Cowsik[8].) Accelerator experiments restrict the mixing angles for still heavier neutrinos, and could lead to very stringent limits.

Our astrophysical limits are based on consideration of heavy neutrinos produced in the Sun and decaying in interplanetary space. Neutrinos of mass up to 14 MeV can be produced in the Sun through the decay ${}^8\mathrm{B} \to {}^8\mathrm{Be} + \mathrm{e}^+ + \nu$ in a side branch of the proton–proton chain. (These neutrinos are those searched for by R. Davis. He finds a factor three less than the lower theoretical estimate. We use the lower theoretical value here but appropriate factors of three may be applied.) If the coupling of the heavy neutrinos is a factor $\sin\theta_e$ of the total, then the expected heavy neutrino flux at Earth will be $F = \Phi \sin^2\theta_e g(m)$ where $2.4\times10^6 < \Phi < 4.5\times10^6\,\mathrm{cm^{-2}\,s^{-1}}$ is the flux calculated for massless neutrinos[9] and $g(m)$ is a factor taking into account the reduced phase space for ${}^8\mathrm{B}$ decay ($g(0) = 1$). The decay rate Γ for heavy neutrinos into electron, positron and light neutrino is obtained by scaling from the μ lifetime as

$$\begin{aligned}\Gamma &= \Gamma_{\mu\to e\nu\nu}(m_\nu/m_\mu)^5 \sin^2\theta_e h(m_e/m_\nu)\\ &= 5\times10^{-5}(m_\nu[\mathrm{MeV}])^5 \sin^2\theta_e h(m_e/m_\nu)\,\mathrm{s}^{-1} \qquad (2)\end{aligned}$$

where h is a phase-space factor with $h(0) = 1$. In the interesting range of parameters the lifetime of the heavy neutrino is so long that most of them decay outside the Earth's orbit. To calculate the flux of positrons at the Earth's orbit we simply multiply the neutrino flux by the probability of decay inside the Earth's orbit (actually this must be done energy by energy to take account of time dilatation).

Now we can compare these calculations with the relevant experiments[10–12]. According to the latter, the flux of positrons in interplanetary space at around 5 MeV kinetic energy is $\simeq 10^{-4}\,\mathrm{cm^{-2}\,s^{-1}\,sr^{-1}\,MeV^{-1}}$. This flux is consistent with estimates of positron flux from cosmic-ray secondaries, so we take it as an upper limit on the possible flux from solar neutrino decay. Comparing with our calculations, we find severe constraints on the mixing angle. For example, at $m_\nu = 2\,\mathrm{MeV}\ \sin^2\theta_e < 8\times10^{-5}$, at $m_\nu = 5\,\mathrm{MeV}\ \sin^2\theta_e < 5\times10^{-6}$, at $m_\nu = 10\,\mathrm{MeV}\ \sin^2\theta_e < 1\times10^{-6}$. The constraints continue to be quite severe until the mass gets very close to the kinematic limits—1.1 MeV for the $\mathrm{e^+e^-}$ decay and 14 MeV for production.

Neutrinos of mass <1.44 MeV would also be produced in the reaction $\mathrm{p + p + e^- \to d} + \nu$. The estimated neutrino flux from this reaction is $1.6\times10^8\,\mathrm{cm^{-2}\,s^{-1}}$, about two orders of magnitude larger than the flux from ${}^8\mathrm{B}\to{}^8\mathrm{Be}$, and because these neutrinos would have low energy, a larger fraction of them would decay inside the Earth's orbit. However, the experimental limit on positrons in this energy range is about two orders of magnitude higher than at 5 MeV, so the limit on $\sin^2\theta_e$ is only slightly better.

Dicus *et al.*[13] have derived constraints based on cosmological considerations which are complementary to ours. Heavy neutrinos which are long-lived are dangerous in cosmology: their decay can inject additional entropy into the Universe after nucleosynthesis, thus invalidating calculations of deuterium abundance; or their decay can introduce non-thermal background radiation contrary to observation; or their failure to decay can cause contraction of the Universe as we mentioned before. The limits of Dicus *et al.* seem to overlap ours, so that if we take both the big-bang cosmology and solar neutrino production seriously, neutrinos in the mass range 1.1–14 MeV seem to be ruled out altogether. We leave aside the bizarre possibility of a heavy neutrino which although above the threshold for $\mathrm{e^+e^-}$ decays only into three neutrino channels.

Accelerator experiments can also put constraints on heavy neutrinos, using the same principles (with the Sun replaced by an accelerator). To get an idea of the magnitudes, let us compare the mean free path l_m for weak interactions of a neutrino with kinetic energy E with the decay length l_d for decay into $\mathrm{e^+e^-}$.

$$\begin{aligned}l_m &\simeq (G^2 m_p E n)^{-1} = \frac{4\times10^{13}}{E[\mathrm{GeV}]}\ \mathrm{cm}\\ l_d &\simeq 6\times10^{17}\frac{E[\mathrm{GeV}]}{(m[\mathrm{MeV}])^6\sin^2\theta_e}\ \mathrm{cm}\end{aligned} \qquad (3)$$

Therefore, neutrinos with masses of 50 MeV or so might be more likely to decay yielding an energetic $\mathrm{e^+e^-}$ pair than to undergo normal weak interactions in a large detector. To make the comparison fair, note that other mixing factors appear in the relative production rates for heavy versus light neutrinos. If the neutrino beam is generated from π decays, a heavy neutrino with mass <35 MeV can come from $\pi^+\to\mu^+\nu$ with a mixing factor $\sin^2\theta_\mu$ measuring its participation in the muon charged current. Another source could be $\pi^+\to\mathrm{e}^+\nu$, with the usual kinematic suppression undone for heavy neutrinos: this would occur with probability $\sin^2\theta_e(m_\nu/m_\mu)^2$ of the ordinary π decay. Similar considerations apply to K-decay which also has larger available phase space and could therefore produce heavier neutrinos.

Note that heavy neutrinos produced in $\pi^+\to\mathrm{e}^+\nu$ decay would be largely helicity-flipped. If the neutrino mass is of Majorana type, then these helicity-flipped neutrinos would interact and decay as antineutrinos. The angular distribution of the $\mathrm{e^+e^-}$ pairs in the decay would be modified accordingly.

We have not examined carefully what quantitative conclusions can be drawn from existing experiments, but clearly heavy neutrinos with mass >50 MeV would have to appear with small mixing angles. It would seem that very sensitive experiments could be designed to look for ν decays in flight, in particular at a K-meson factory.

The effect of such a massive neutrino on the e/μ ratio in K_{l2} and π_{l2} decays has been calculated by Goldman and Kolb[14]. Recently Shrock[15] has proposed searching for unstable neutrinos or for characteristic energy shifts of the charged leptons in π_{l2} and K_{l2} decays, and Calaprice (S. Treiman, personal communication) has analysed deviations from universality in nuclear decays as possible signatures for large neutrino masses. These methods, of course, could in principle be sensitive to masses below 1.1 MeV.

We thank S. Treiman, S. Peale, E. Hones, R. Lingenfelter and E. Stone for useful conversations. This work was supported by NSF grant PHY77-27084.

Received 27 October; accepted 15 December 1980.

1. Lyubimov, V. A., Novikov, E. G., Nozik, V. Z., Tretyakov, E. F. & Kosik, V. S. *Phys. Lett.* **94B**, 266 (1980).
2. Reines, F., Sobel, H. W. & Pasierb, E. *Phys. Rev. Lett.* **45,** 1307 (1980).
3. Witten, E. Preprint HUTP-80/A031 (Harvard University, 1980).
4. Faber, S. & Gallagher, J. *A. Rev. Astr. Astrophys.* (1979).
5. Cowsik, R. & McLelland, J. *Phys. Rev. Lett.* **29,** 669 (1972).
6. Tremaine, S., Gunn, J. *Phys. Rev. Lett.* **42,** 407 (1979).
7. Gunn, J., Lee, B., Lerche, I., Schramm, D. & Steigman, G. *Astrophys. J.* **223,** 1015 (1978).
8. Cowsik, R. *Phys. Rev. Lett.* **39,** 784 (1977).
9. Bachall, J. & Ulrich, R. *Astrophys. J. Lett.* **160,** L57 (1970).
10. Beurmann, K., Rice, C. Stone, E. & Vogt, R. *Acta phys. hung.* **29,** Suppl. 1, 173 (1970).
11. Cline, T. & Porreca, G. *Acta phys. hung.* **29,** Suppl. 1, 145 (1970).
12. Mewaldt, R., Stone, E. & Vogt, R. *14th int. Cosmic Ray Conf.* **1,** 349 (1975).
13. Dicus, D. Kolb, E., Teplitz, V. & Wagoner, R. *Phys. Rev.* **D17,** 1529 (1978).
14. Kolb, E. W. & Goldman, T. *Phys. Rev. Lett.* **43,** 897 (1979).
15. Shrock, R. *Phys. Lett.* **96B,** 159 (1980); Stonybrook preprint ITP-8SB-80-56 (1980).

IR observations of the solar corona—a ring around the Sun?

U. R. Rao, T. K. Alex, V. S. Iyengar,
K. Kasturirangan, T. M. K. Marar, R. S. Mathur
& D. P. Sharma

ISRO Satellite Centre, Bangalore 560 058, India

The possibility of a permanent ring structure around the Sun was first pointed out by Brecher *et al.*[1] from evolutionary, physical and chemical considerations as well as constraints arising from observational astronomy related to Sun and its immediate neighbourhood. In particular, Brecher *et al.*[1] concluded that such a ring, located at ~4 $R_\odot$, must consist of refractory particles (graphite, for example) of size >10 km at ~1,000–2,000 K which, therefore, would emit thermal radiation peaking around 1.5–3.0 µm. We report here our attempt to detect the existence of this ring structure utilizing the IR observations taken during optimal viewing conditions of the total solar eclipse of 16 February 1980.

The experimental system consisted of a 6-inch $f/6$ newtonian telescope coupled to a celostat. The detector assembly consisted of a PbS detector, a chopper with its motor, a filter and a detector cooling box. The 1×1 mm PbS detector, cooled to liquid nitrogen temperatures had an NEP of 1.3×10^{-10} W at a chopping frequency of 1.042 kHZ. The interior of the cooling box was continuously purged by dry argon gas to prevent frosting of the filter–detector assembly. The IR filter with antireflection coating had a transmission of ~75% in the 1.9–2.8 µm range. The electronics consisted of a preamplifier (ORTEC model 5001), lock-in amplifier (ORTEC model 9501), chopper reference circuit and a tape recorder. The entire detector assembly was mounted on a moving carriage that enabled bi-directional scanning of the corona in the image plane, up to 5.5 $R_\odot$ on either side of the solar disk. The instrument was calibrated using the Moon at its various phases, and a black-body source.

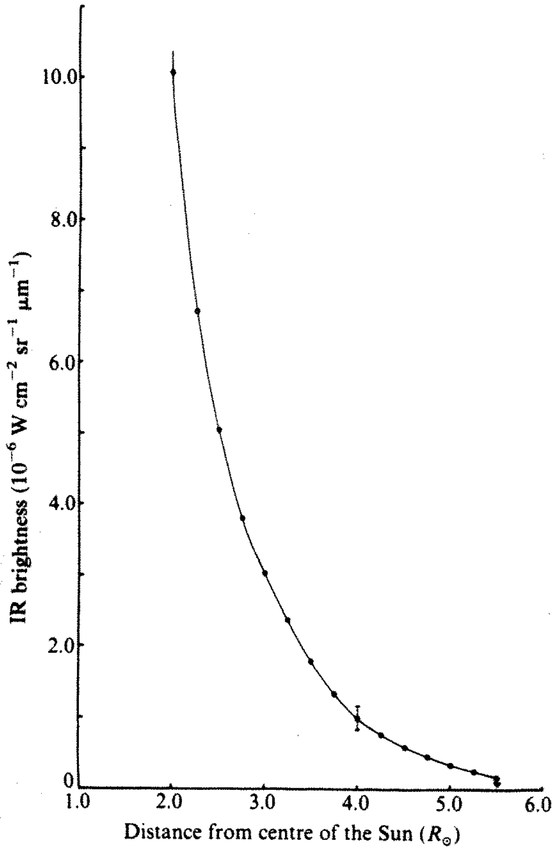


Fig. 1 The IR brightness of the solar corona in the equatorial plane from 2 to 5.5 $R_\odot$.

The observations were made from Hosur (long 75°09′04″; lat 15°00′12″) in Karnataka, India where the duration of the eclipse totality was ~2.5 min. During totality five scans in the image plane were carried out along 9.5°, 5.5°, 1.5°, −2.5° and −6.5° with respect to the solar equatorial plane. The spatial resolution of the instrument was 0.25 $R_\odot$. A continuous scan was made in the image plane.

Figure 1 shows the measured IR flux plotted as a function of distance from the centre of the Sun to a distance of 5.5 $R_\odot$, corresponding to the scan mode close to the equatorial plane. The absolute brightness of the outer corona has been deduced using calibration data on the Moon at its various phases. Comparison of the flux obtained in the present experiment at 3 $R_\odot$ with values obtained by Peterson[2] and MacQueen[3] indicate that the present value is larger by ~3.5 and 6 times their values respectively. These differences can be reconciled, as the present observations coincided with the period of solar maximum and the earlier observations of Peterson[2] and MacQueen[3] were made during the total solar eclipse of November 1966 at a time of solar minimum. Three independent observations[4] also indicate the variabilities in the IR emission from the solar corona. The first observation, using a coronagraph during a 1971 balloon flight, apparently revealed a weak and diffuse emission around 4.3 $R_\odot$ at 2.2 µm. The second observation, during a 1973 total solar eclipse revealed no evidence of a feature at 4 $R_\odot$ nor of a shelf of emission out to 6 $R_\odot$. In the third observation, a balloon flight after the 1973 eclipse also showed a weak feature, this time at 5 $R_\odot$ elongation, but no shelf. Our observations also imply that the IR corona is highly variable. Continuous monitoring of the solar corona in the near IR is, therefore, important for understanding interplanetary dust models.

At 4 $R_\odot$, the expected location of the ring, our measured flux is a factor of only 1.6 larger than the value obtained by Peterson[2]. We did not, however, observe the excess flux above the continuum reported by Peterson[2] and MacQueen[3]. The present observations therefore do not support the existence of a ring in the equatorial plane of the Sun, corresponding to a 3σ sensitivity of 10^{-7} W cm^{-2} sr^{-1} µm^{-1} at 2.2 µm in our detection system. Other scans which were conducted away from the equatorial plane have shown similar results, namely the absence of IR features above the level of detectivity. Thus our observations place definitive constraints on the hypothesis of Brecher *et al.*:

(1) The ring, if it exists in the equatorial plane between 2 and 5.5 $R_\odot$, has an IR brightness lower than 10^{-7} W cm^{-2} sr^{-1} µm^{-1} at 2.2 µm. The excess emission observed at 4 $R_\odot$ by Peterson and MacQueen cannot be ascribed to a 'stable' ring structure, because if the dispersion time for such a ring were of the order of 10^{13} yr we should have seen it in the present experiment.

(2) The present upper limit would also imply an upper limit to the total number of 10 km objects of ~10^8, based on the visual magnitude $m_v = +8$ to $+10$ for the individual objects as hypothesized by Brecher *et al.*

(3) If the ring resides at $R \gtrsim 5.5\ R_\odot$, the total mass in the ring is $\lesssim 3\times10^{25}$ g which in turn will limit the total number of 10 km objects (made of graphite with a density of 2 g cm^{-3}) to ~3×10^6. If the ring consisted of less refractory materials and were located at distances greater than 5.5 $R_\odot$ from the Sun, our experiment would not have detected it. If smaller particles with size <10 km, that are spiralling into the Sun, contribute significantly to the total IR emission of the solar corona at distances less than 5.5 $R_\odot$, it is possible that such radiation may mask the excess emission arising from a ring of larger particles with a sharp outer boundary within 5.5 $R_\odot$. But then the observation of an excess emission at times, makes the problem more enigmatic and interesting for future measurements.

0028–0836/81/090779—02$01.00

If a ring structure exists around the Sun between 2 and 5.5 $R_\odot$ in the equatorial plane, this experiment places an upper limit of 10^{-7} W cm^{-2} sr^{-1} μm^{-1} for its emission at 2.2 μm. Further, observations seem to indicate a high degree of variability in the near IR emission from the solar corona.

We thank the staff from the ISRO Satellite Centre for their cooperation, and the Department of Space, Government of India for Financial support.

Received 30 June; accepted 26 November 1980.

1. Brecher, K., Brecher, A., Morrison, P. & Wasserman, I. *Nature* **282,** 50–52 (1979).
2. Peterson, A. W. *Astrophys. J. Lett.* **148,** L37–L39 (1967).
3. MacQueen, R. M. *Astrophys. J.* **154,** 1059–1076 (1968).
4. Strong, J. *NASA TMX 62397*, Pap. 1 (1974).

Temperature dependence for the power outputs of *n*-CdSe liquid junction cells

J. F. McCann, M. Skyllas Kazacos & D. Haneman

School of Physics, The University of New South Wales, PO Box 1,
Kensington, New South Wales 2033, Australia

Thin film *n*-CdSe photoanodes have great potential as efficient and inexpensive electrodes in liquid junction photovoltaics[1–3]. This stems from the stability of *n*-CdSe in the S/S^{2-} electrolyte, its direct band gap[4] of 1.7 eV and the large barrier height, $\Phi_B(T)$, of ~1 eV at the *n*-CdSe–S/S^{2-} interface[5]. The recently reported 8% AM1 conversion efficiency for a cadmium chalcogenide thin film based cell[6] suggests that these low cost devices may soon compete favourably with the conversion efficiencies of the more expensive single crystal solid state photovoltaics. In view of the increasing practical importance of this system, the study of any parameter which may affect its performance is relevant. We report here the dramatic effect of temperature on the power outputs of *n*-CdSe based cells.

Single crystals of *n*-CdSe (Cleveland Crystals Inc.) were fabricated into electrodes by forming an ohmic contact on the back side of each crystal with In:Hg amalgam and soldering onto a copper wire. The crystals were mounted in epoxy resin, polished and etched by the usual techniques[1] before use. Thin films of *n*-CdSe were prepared on 1×3 cm^2 Ti strips by the electroreduction of Na_2SeSO_3 in the presence of a high concentration of Cd^{2+} (ref. 7). The films were annealed in a vacuum at 400 °C for 1 h. Several CoS counter electrodes were prepared by the method described by Hodes *et al.*[8]. The electrolyte used in this study consisted of S(1M)–Na_2S(1M)–Se(0.07M)–NaOH(1M) and was prepared with AR grade reagents and distilled water. During all experiments the electrolyte was continuously purged with N_2 and stirred. Cell current–voltage (*i–V*) curves were obtained by varying a load resistor between the *n*-CdSe and CoS electrodes. A 300-W General Electric ELH lamp was used to illuminate the semiconductor electrodes by a Schott KG-2 filter (intensity ~ 70 mW cm^{-2}).

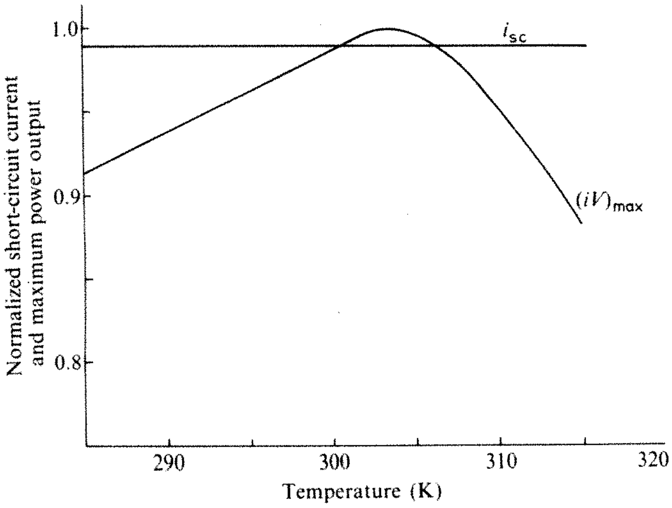


Fig. 1 The normalized short-circuit current, i_{sc}, and maximum power output, $(iV)_{max}$ of single crystal *n*-CdSe/S(1M)–Na_2S(1M)NaOH(1M)–Se(0.07M)/CoS as a function of temperature.

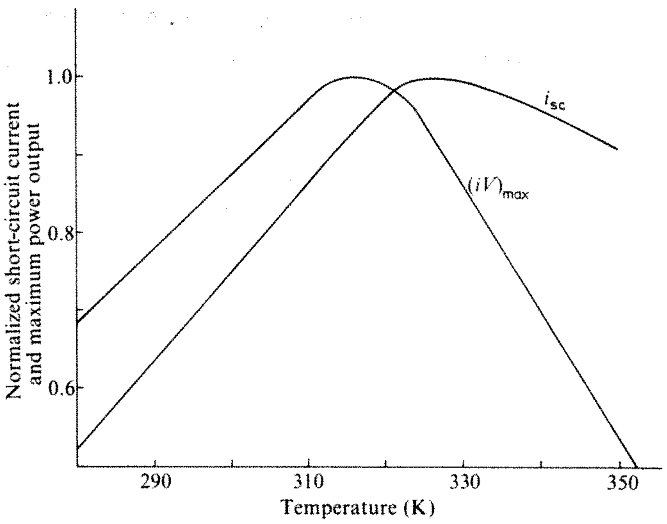


Fig. 2 The normalized short-circuit current and maximum power output of a thin film *n*-CdSe/S(1M)–Na_2S(1M)–NaOH(1M)–Se(0.7M)/CoS cell as a function of temperature.

The resistance of a cell with the electrodes separated by ~1 cm was measured by passing a sinusoidal voltage (Heathkit ET-3100 sine wave generator) of the desired frequency across two resistors (R_1 and R_2) which were in series with each other and with the cell. A PAR lock-in amplifier (model 124A) was used to measure the in phase signal, Z_R, differentially across R_2 (typically 1 Ω) and directly across the cell (resistance R_{cell}) which was kept at open circuit. The amplitude of the sine wave across the cell was kept <10 mV by adjusting the value of R_1. The value of R_{cell} was determined by comparing the voltage reading across R_2 with the voltage reading across the cell.

Series of *i–V* curves were determined for several *n*-CdSe cells over a range of temperatures. These were obtained by first cooling down the electrolyte in the single compartment cell, and gradually heating it on a heating mantle over a period of ~2 h. *i–V* curves were obtained at various temperature intervals and the electrodes were illuminated only while the curves were being recorded. To verify that the effects observed were due to temperature and not time, the solution was then allowed to cool down and *i–V* curves again recorded. The data obtained on cooling were equivalent to those obtained on the heating cycle. The open circuit voltages of both single crystal and thin film based cells dropped by ~2 mV per degree C as the temperature was increased, which is comparable with the decrease in V_{oc} typically observed in solid state photovoltaics for comparable illumination intensities. In the latter systems, this trend can be predicted from the relation for the open circuit voltage, $V_{oc}(T)$ as a function of temperature[5]:

$$V_{oc}(T)=\frac{nkT}{q}\ln\left[\frac{i_{ph}}{A^*T^2\exp(-\Phi_B/nkT)}+1\right] \qquad (1)$$

where q is the electronic charge, k is Boltzmann's constant, T is the absolute temperature, n is a factor which takes the non-ideal behaviour of the semiconductor into account, i_{ph} is the semiconductor photocurrent, and A^* is Richardson's constant for a given semiconductor. In the case of a photoelectrochemical cell, however, this relationship has to be modified to include the

0028–0836/81/090780—03$01.00

temperature dependences of the redox potential of the electrolyte, and the flat band potential of the semiconductor. These factors could result in a temperature dependence of the barrier height Φ_B. Variations in the semiconductor band gap and solution absorption with temperature could also have an effect on i_{ph}. The A^*T^2 term in equation (1) would also need to be modified to take the electron transfer probability through the Helmholtz layer into the electrolyte into consideration. In spite of these limitations of equation (1) when applied to a liquid junction photocell, the observed decrease in V_{oc} for n-CdSe in S/S^{2-} of 2 mV per degree C for increasing temperature is in agreement with the value predicted by this relation and could suggest that the modifications outlined above may not be very significant.

The typical effects of temperature on the short-circuit current and power outputs of both single crystal and thin film n-CdSe cells are shown in Figs 1 and 2 respectively. Although the short circuit current for the single crystal cell remains essentially constant over the temperature range, the power output of the cell exhibits a maximum at ~30 °C. This is due to a corresponding increase in the fill factor with temperature which counteracts the drop in V_{oc} and thus results in the maximum observed for the power output of the cell. The behaviour of the thin film cell shows a more dramatic temperature dependence. In contrast to the single crystal cells, the short-circuit current for this type of cell increases drastically up to about 50 °C. Experiments on several thin film electrodes have indicated that the short circuit current tends to level off beyond 50 °C rather than drop slightly as seen in Fig. 2. There is a corresponding improvement in the output power up to ~40 °C. However, due to the continuing decrease in V_{oc} and the levelling off of the short circuit current, the power output decreases rapidly at higher temperatures resulting in an optimum cell operating temperature of around 40 °C. We suggest that the striking effect of temperature on the short circuit current for thin film cells can be explained in terms of the dissolution kinetics of sulphur formed at the photoanode. Because of the polycrystalline nature of these films the electrolyte penetrates the fine pores. As the oxidation of sulphide progresses, the sulphide ion concentration within the pores becomes depleted. Hence the dissolution of sulphur formed within the pores, which occurs according to the reaction:

$$\mathrm{S(solid)} + \mathrm{S}_x^{2-} \rightleftarrows \mathrm{S}_{x+1}^{2-} \text{ (in solution)} \qquad (2)$$

is limited by the rate of diffusion of sulphide ions into the pores from the bulk of the solution and partial passivation of the electrode occurs. The rate of sulphur dissolution increases with temperature until it is no longer a limiting factor. This causes the corresponding increase in the short circuit current observed up to ~50 °C (see Fig. 2). Films grown by another technique (D. Miller and D.H., unpublished results) also show the temperature behaviour described above.

As a further test of the above hypothesis, experiments were carried out over a range of light intensities on a CdSe film grown by the latter technique. The kinetics of sulphur dissolution would be expected to have a much less important effect at comparatively low light intensities. Measurements between 0 and 40 °C showed that the rate of increase of the short circuit current with temperature dropped from 0.33 at a light intensity of ~140 mW cm^{-2} to 0.15 at 70 mW cm^{-2} to 0 at 0.5 mW cm^{-2}. This behaviour is therefore consistent with the sulphur dissolution hypothesis.

The effect of temperature on the electrolyte resistance was also examined as this may also influence the i–V characteristics of n-CdSe cells. Figure 3a shows the resistance of a n-CdSe/S–S^{2-}–Se–OH^-/CoS cell plotted as a function of frequency both in the dark and under illumination, and shows that the resistance levels off at frequencies ≥20 kHz. This graph also illustrates the decrease in the net resistance of the semiconductor on illumination which is apparently due to a reduction in the resistance of the space charge region and an increase in the photoconductivity of the semiconductor bulk.

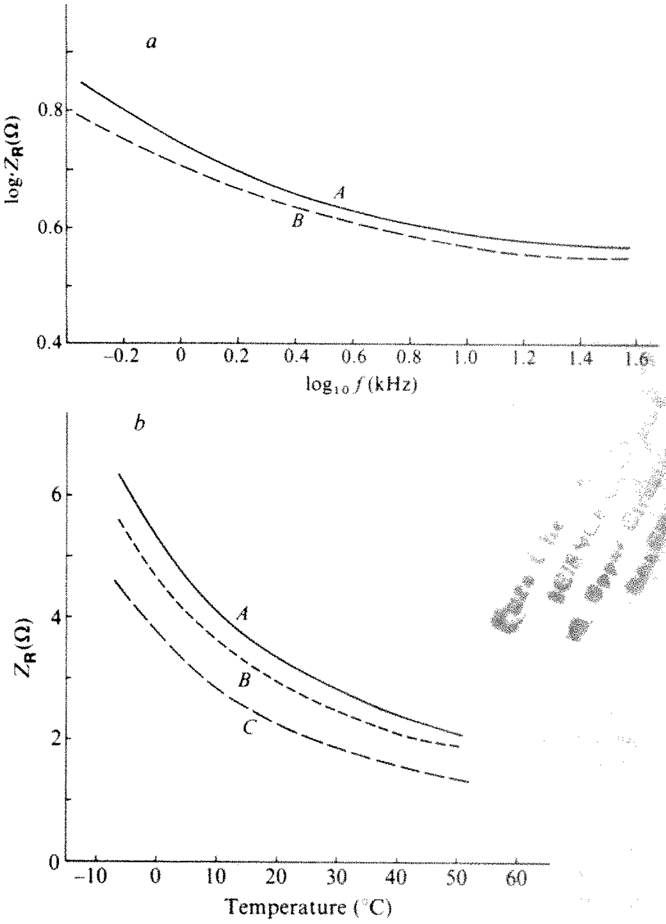


Fig. 3 a, The real component of the impedance, Z_R, of a thin film n-CdSe/S(1M)–Na_2S(1M)–NaOH(1M)–Se(0.7M)/CoS cell plotted against frequency f (in kHz): A in the dark and B under illumination; T = 22.5 °C. b, Z_R at 28 kHz versus temperature for a thin film n-CdSe/S(1M)–Na_2S(1M)–NaOH(1M)–Se(0.7M)/CoS cell: A in the dark and B under illumination. C, Pt–Pt cell in the same electrolyte.

Using an input frequency of 28 kHz, the above cell resistance was then measured as a function of temperature and compared with the results obtained using two platinum electrodes in the S/S^{2-} electrolyte. These results are presented in Fig. 3b. Again, the decrease in cell resistance on illumination of the CdSe can be observed in curves A and B. These curves also show the significant drop in cell resistance with increasing temperature. The parallel trend obtained with the two platinum electrodes (curve C) indicates that the temperature dependence is predominantly due to the electrolyte resistance rather than the semiconductor resistance varying as a function of temperature.

The above results therefore suggest that the changing resistance of the electrolyte as a function of temperature would be expected to influence the shapes of the cell i–V curves obtained for liquid junction solar cells. In fact, for the illuminated CdSe thin film solar cell used here, this resistance is seen to change from ~4.6 Ω at 0 °C to ~2 Ω at 50 °C. A significant increase in the fill factor of the cell i–V curves would be expected over this temperature range and was, in fact, observed in a previous study[5]. Another factor which may also influence the fill factor as a function of temperature is the change in the polarization of the CoS counter electrode with temperature. This latter effect was not examined in the present study but the Weizmann group have shown that polarization at a CoS counter electrode decreases with temperature[10]. This would also improve the fill factor of a cell with temperature.

The above results emphasize the importance of controlling the operating temperatures of n-CdSe liquid junction cells particularly those based on thin films to obtain the maximum possible photo-conversion.

This work was supported by the National Energy Research Development and Demonstration Council. M. Skyllas Kazacos has been supported by a Queen Elizabeth II Fellowship.

Received 7 July; accepted 10 December 1980.

1. Miller, B. *et al. J. electrochem. Soc.* **124,** 1919 (1977).
2. Hodes, G., Manassen J. & Cahen, D. *Nature* **261,** 403 (1976).
3. Owen, J. R. *Nature* **267,** 504 (1977).
4. Williams, R. *J. chem. Phys.* **32,** 1505 (1960).
5. McCann, J. F., Kazacos, Skyllas M. & Haneman, D. *J. electrochem. Soc.* (submitted).
6. Cahen, D., Hodes, G. & Manassen, J. *Abst. 179th ACS Natn Meet.*, Houston (ACS, Washington, 1980).
7. Skyllas Kazacos, M. & Miller, B. Patent (filed).
8. Hodes, G., Manassen, J. & Cahen, D. *J. appl. Electrochem.* **7,** 181 (1971).
9. Tomkiewicz, M. *J. electrochem. Soc.* **126,** 2220 (1979).
10. Hodes, G., Manassen, J. & Cahen, D. *J. electrochem. Soc.* **127,** 544 (1980).

Zoned aluminium distribution in synthetic zeolite ZSM-5

R. von Ballmoos & W. M. Meier

Institut für Kristallographie und Petrographie, ETH Zürich, 8092 Zürich, Switzerland

ZSM-5, a prominent member of the new generation of high silica zeolites, is a potential catalyst. The activity of ZSM-5 type synthetic zeolites with respect to acid-catalysed reactions is a linear function of the Al content[1] which can assume extremely low values, but in catalytic applications is typically 1–3 Al per unit cell containing 96 Si + Al atoms. We now present substantial evidence that the Al distribution in ZSM-5 crystals is non-uniform. We first noted zoning of Al in ZSM-5 when examining scanning X-ray micrographs (Al Kα) of 50 μm crystals. The comparatively large crystals necessary for detailed measurements of Al concentration profiles using electron probe analysis suggest that Al is invariably concentrated in the rim portion of the crystals. No comparable zoning has been reported previously for zeolites.

ZSM-5 is synthesized in siliceous systems containing tetrapropylammonium (TPA) and sodium ions[2]. The Si/Al ratio of ZSM-5 ranges from ~10 to >4,000 (ref. 1). Such figures refer to bulk composition of the crystals which typically do not exceed a few micrometres in size. The crystal structure of ZSM-5 (ref. 3) and the structural properties of its aluminosilicate framework have been described elsewhere[4–6]. X-ray structure analysis could not be expected to provide information on the arrangement of the relatively few Al atoms and a uniform distribution has been assumed.

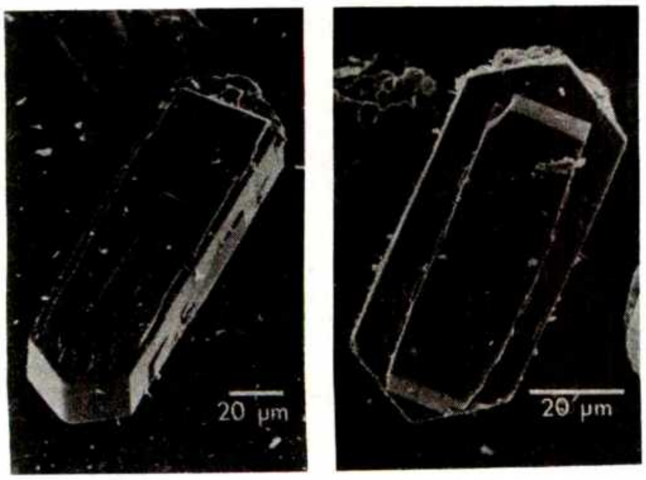


Fig. 1 Scanning electron micrographs of typical large ZSM-5 crystals. Gel particles are visible at the upper left.

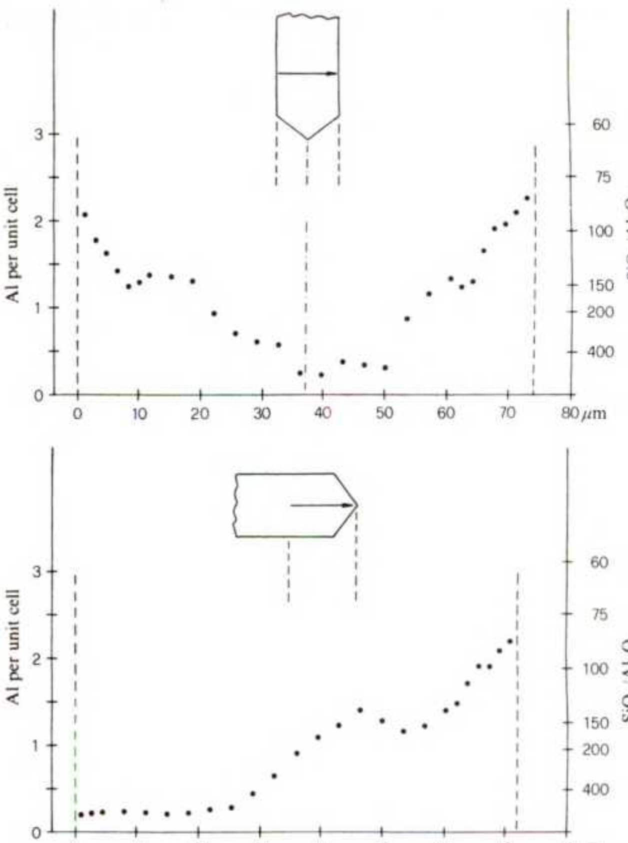


Fig. 2 Two perpendicular profiles of Al concentration across a section of a ZSM-5 crystal determined by microprobe step scans (shown in the inserted outline).

Appropriate conditions for growing large ZSM-5 crystals were determined, using dilute systems containing ~40–60 mg of solid per 10 cm^3 of solution. A suitable molar composition is around $11SiO_2:Al(NO_3)_3:2NaOH:72(TPA)OH:3{,}300\ H_2O$. Crystallizations were carried out in small steel autoclaves (25 cm^3) at 200 °C for 5 days[7]. The largest crystals obtained were 60–200 μm. Figure 1 shows scanning electron micrographs of typical crystals which appeared invariably twinned. All solid products of the crystallizations contained a considerable amount of gel. The Si/Al ratio of ~3–4 for the gel contrasts strongly with that of the crystals with bulk Si/Al ratios of 50–100. X-ray powder diffraction patterns were used to identify the ZSM-5 crystals.

Five specimens were filled into holes of 4 mm diameter in an acrylate disk. The holes were then filled with an epoxy resin (Epotek 330) which was then cured overnight at 60 °C. Grinding was completed with 0–3 μm diamond paste on a lapping wheel. The sample disks were cleaned with hexane and stored in water-free atmosphere. All samples were beryllium coated, and for some a very thin film of aluminium was first applied to assure better contact of the beryllium coating. Crystals of the silica end member of the ZSM-5 compositional series ('silicalite'[8]) were also prepared to monitor possible aluminium contamination. Electron microprobe scans used an ARL-EMX instrument at 15 kV, 0.08 μA sample current, a beam diameter of 0.8 μm producing an analysis volume of about 2 μm^3 and an average counting time of 8 s for each point, with crystal spectrometers (ADP and RAP). Orthoclase was used as a standard and the computer programs EMX and EMMA were used to process the data. Step advancement was mechanical at intervals of 1.8 or 3.5 μm along a straight path. Only true cross-sections through the core of the crystal and oriented approximately parallel or perpendicular to *c* were analysed. Scans were taken crosswise for better centering.

A total of 60 scans were made of 40 large ZSM-5 crystals with representative Si/Al bulk ratios of 50–100. All Al concentration profiles showed pronounced zoning as in Fig. 2. The Al concentration in the core was invariably below that in the rim. Minimum and maximum concentrations of Al within a ZSM-5 crystal frequently differ by a factor of 5 or more. The most pronounced zoning, with Al concentration in the rim exceeding that in the core by ≫10, was observed in Fig. 3. Core concentrations were often found near 0.2 Al per unit cell. In some crystals the Al concentration in the core was not above a detection level of 0.05 Al atoms per unit cell. Usually zoning is monotonic from a siliceous core to an Al-rich rim but some waviness (Figs 2, 3) could frequently be observed. (More details are presented elsewhere[7].) All zoning was essentially symmetric.

Strong zoning of this kind has not so far been found in zeolites. There is a relative increase in the Si/Al ratios from core to rim of large NaX crystals of a few per cent[9]. This zoning in NaX is reversed and much less striking than that in ZSM-5. Crystals of natural faujasite, phillipsite and offretite were found to be uniform[10]. On the other hand, chemical zoning is present in feldspars of the plagioclase series[11].

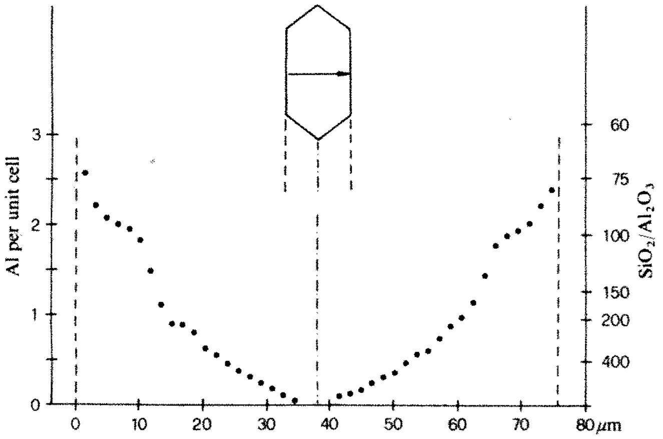


Fig. 3 Example of a scan which exhibits Al-zoning with extreme concentrations differing by a factor of more than 10. The innermost core concentration is around the detection limit of 0.05 Al per unit cell.

The presence of Al-rich solid gel in the products might suggest that the increased Al content in the rim is associated with dissolution of gel particles which enhances the Al content of the solution towards the end of crystallization. This would imply crystal growth from the liquid phase.

Although the zoning observed in large ZSM-5 crystals does not necessarily apply to much smaller crystals, their bulk compositions should be interpreted with caution.

We thank R. Gubser (microprobe), E. Schaerli (sample preparation) and P. Waegli (SEM) for technical assistance, and Dr D. H. Olson and others at Mobil Central Research Laboratories for helpful discussions. This study was supported by Swiss NSF grant 2.114-0.78.

Received 3 November; accepted 8 December 1980.

1. Olson, D. H., Haag, W. O. & Lago, R. M. *J. Catal.* **61,** 390–396 (1980).
2. Argauer, R. J. & Landolt, G. R. *U.S. Patent* **3,** 702, 886 (1972).
3. Kokotailo, G. T., Lawton, S. L., Olson, D. H. & Meier, W. M. *Nature* **272,** 437–438 (1978).
4. Kokotailo, G. T. & Meier, W. M. *Chem. Soc. Spec. Publ.* **33,** 133–139 (1980).
5. Wu, E. L. *et al. J. phys. Chem.* **83,** 2777–2781 (1979).
6. Olson, D. H., Kokotailo, G. T., Lawton, S. L. & Meier, W. M. *J. phys. Chem.* (submitted).
7. von Ballmoos, R. thesis, ETH Zurich (1981).
8. Flanigen, E. M. *et al. Nature* **271,** 512–516 (1978).
9. Weeks, T. J. & Passoja, D. E. *Clays Clay Miner.* **25,** 211–213 (1977).
10. Rinaldi, R., Smith, J. V. & Jung, G. *Neuses Jb. Miner. Mh.* 433–443 (1975).
11. Smith, J. V. *Feldspar Minerals* Vol. 2, 206–247 (Springer, Berlin, 1974).

0028-0836/81/090783—03$01.00

Volcanism and the dynamics of open magma chambers

Stephen Blake

Lunar and Planetary Unit, Department of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ, UK

The repeated eruption of many volcanoes is now accounted for by events taking place in a subterranean magma chamber, which is 'open' in the sense that the inflow of material to the chamber from below leads to a magma pressure P which exceeds the lithostatic pressure P_L by an amount greater than the tensile strength of the chamber walls (σ) before the contents of the magma chamber have reached thermodynamic and geochemical equilibrium. This is the condition which must be satisfied when an eruption occurs. We have now derived a relationship giving the critical volume of added material (ΔV) required to cause an eruption as a function of magma compressibility and the volume by which the chamber expands.

Geochemical, petrological and stratigraphic studies of lava piles at individual volcanoes have demonstrated than many sub-volcanic magma chambers are open systems where an evolutionary trend controlled by crystal fractionation is modified by the episodic influx of less evolved magma[1–4]. Disequilibrium phenocrysts and glasses in some lavas and ejecta also show that magma mixing has preceded eruption[5–10]. Such an open system of replenishment and eruption extends the lifetime of a chamber by buffering heat loss and is capable of chemically[11] and thermally[12] determining the evolution of the magma, depending on the fluid properties and dynamics of the magmas involved[13,14]. In addition to their petrological significance, open magma chambers are important volcanologically. The inflation of volcanoes, often accompanied by seismic activity, is attributable to the influx of magma into a sub-volcanic storage system. Eruption begins when some critical amount of magma has accumulated causing the reservoir to burst[15–19] and deflation typically accompanies withdrawal or eruption of magma[20].

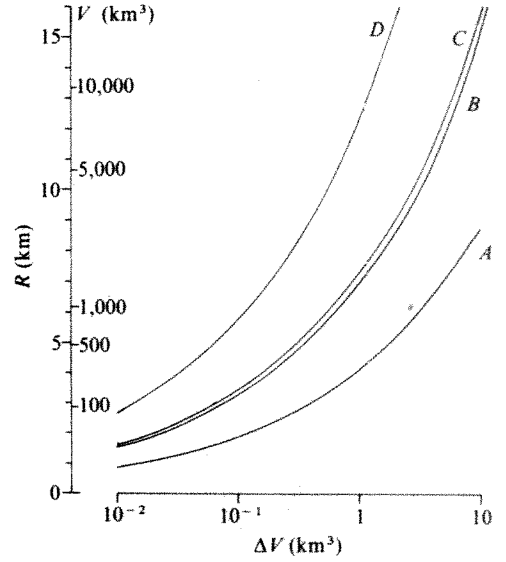


Fig. 1 The volume of magma, ΔV, which must be introduced into a chamber of constant volume, V_c, to trigger eruption according to $\Delta V/V_c = \sigma/b$. Also plotted is the radius, R, of a spherical chamber having volume V_c, but other geometries may be considered. Curves A and B, olivine basalt $b = 14$ GPa; C and D, andesite $b = 80$ GPa; A and C, strong volcano $\sigma = 50$ MPa; B and D, weak volcano $\sigma = 10$ MPa. For each case the critical values of σ/b are: A, 3.6×10^{-3}; B, 6.25×10^{-4}; C, 7×10^{-4}; D, 1.25×10^{-4}.

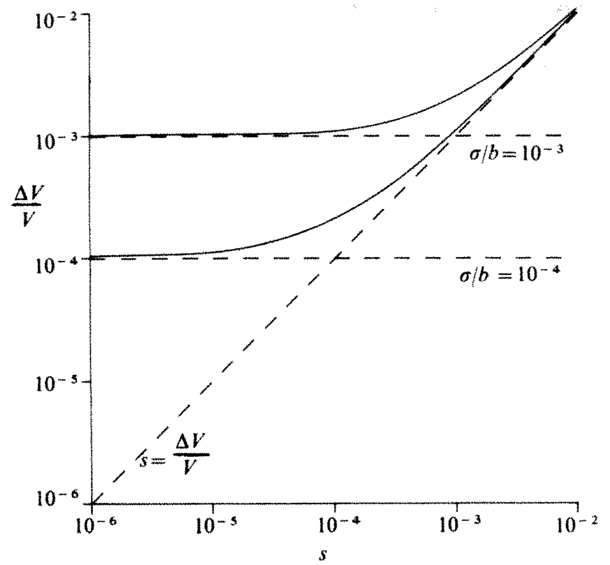


Fig. 2 Allowing a chamber to expand to a size $V_c\,(1+s)$ before eruption causes the critical ratio $\Delta V/V_c$ to increase according to $\Delta V/V_c=\sigma/b(1+s)+s$, which is shown by the heavy lines for the cases σ/b equal to 10^{-3} and 10^{-4}. When $s\ll\frac{\sigma}{b}$, $\frac{\Delta V}{V_c}\rightarrow\frac{\sigma}{b}$, approximately the case of a constant volume chamber; when $s\gg\frac{\sigma}{b}$, $\frac{\Delta V}{V_c}\rightarrow s$ and is independent of σ and b.

First, consider a constant-volume magma chamber which in equilibrium conditions (lithostatic pressure) contains a volume of magma V_c. When additional magma of volume ΔV is introduced, the contents of the chamber must become compressed to accommodate the new magma (an assumption which will be justified later). The compressed magma will occupy the volume V_c at a pressure P ($>P_L$), placing the chamber walls under tension. P and ΔV are related by the constitutive equation of the magma

$$V_c=(V_c+\Delta V)\exp\left[-(P-P_L)/b\right]$$

where b is the bulk modulus of the magma and ΔV is measured at P_L. The critical condition is arrived at when $P-P_L=\sigma$ which, for most rocks, lies in the range 10–100 MPa (ref. 21). The strength of volcanoes is, however, reduced by structural defects resulting from fragmental layers, vesicular and jointed lava flows, dyke and sill complexes and a generally stratified structure. Previous workers[19,22,23] have used values for the tensile strength of volcanoes in the range 10–50 MPa. Values of b have been calculated from the seismic wave velocities and densities of rock melts measured by Murase and McBirney[24] and are found to vary from 14 GPa (Galápagos Islands olivine basalt, 1,200 °C) to 80 GPa (Mount Hood andesite, 1,000 °C). The relative magnitudes of σ and b (that is, $\sigma/b\sim10^{-3}$) simplify the critical condition required for failure of the chamber walls to

$$\frac{\Delta V}{V_c}=\frac{\sigma}{b}$$

When the walls have failed, escape of magma is driven by decompression until magmatic pressure within the chamber has returned to lithostatic and only the stable volume of magma V_c is present. The expelled volume, V_e, is therefore equal to ΔV, as decompressional expansion on eruption is negligible.

Given the bulk modulus of a magma (which varies with composition and temperature) and the tensile strength of the volcano of interest, the erupted volume, being approximately equal to the volume required to trigger an eruption, is proportional to the size of the associated magma chamber.

Figure 1 depicts the relationship between ΔV required to trigger eruption from a spherical chamber of radius R and volume V_c. For basaltic volcanoes (Fig. 1, curves A, B) with $0.01\leqslant V_e<1$ km^3, chamber radii of 1–7 km are predicted, in good agreement with the sizes of exhumed gabbroic plutons[25], the seismically determined sizes of basic magma bodies in the upper crust[26] and the areal dimensions of calderas at stratovolcanoes. For a given chamber size and strength, the larger bulk modulus of cooler andesitic magma causes ΔV to be less than that required by a basaltic magma (Fig. 1, curves C, D). The range of V_e for both magma types is essentially the same, however, which therefore implies larger chambers at andesitic volcanoes ($R<12.5$ km) than at basaltic ones, a conclusion which seems to be substantiated, as dioritic and tonalitic plutons are often larger ($R<15$ km) than gabbroic plutons ($R<5$ km).

The assumption that V_c remains constant may not always be appropriate. If s is defined as the fractional volumetric expansion of the chamber, the critical condition for eruption becomes

$$\frac{\Delta V}{V_c}\simeq\frac{\sigma}{b}(1+s)+s$$

Consequently, the critical ratio $\Delta V/V_c$ is dependent on the relative magnitudes of s and σ/b (Fig. 2). The amount of expansion which a given chamber experiences will depend on its host environment. In unfractured competent country rock, dilation of the chamber results from elastic deformation of the wall rocks. As the compressibility of solid rock is up to an order of magnitude less than that of silicate melts[24] the chamber volume will not be significantly increased and $s<\sigma/b$. Accordingly, newly introduced magma is most easily accommodated not by an increase in chamber volume, but by compression of the magma itself. In such cases the constant volume model as discussed above is a valid simplification and $\Delta V/V_c\simeq\sigma/b\simeq10^{-3}$ adequately describes the critical condition for these eruptions (Fig. 1).

High degrees of chamber enlargement ($s>\sigma/b$) are thought to be due to the creation of new reservoir space by crack opening, filling and propagation such as in the rift zones of the Kilauea volcano[27,28]. Conditions at Kilauea and many other mature volcanoes are favourable for this type of behaviour because the magma chamber resides in a readily fissured lava pile. During these intrusive episodes, surface deformation is manifest (refs 15, 29–31).

The effect on the critical condition of allowing chamber expansion (creation of additional storage space) is shown in Fig. 3 for the case $\sigma/b=10^{-3}$ and $s=0$, 10^{-4}, 10^{-3} and 10^{-2}. When s exceeds σ/b, the factor controlling the onset of eruption (the value of ΔV) becomes the new size of the reservoir and is largely independent of σ/b. In these cases, the use of V_e and the

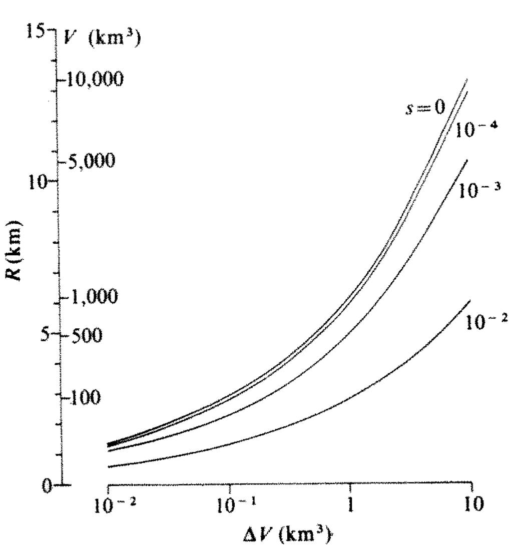


Fig. 3 The effect on ΔV of permitting chamber expansion $s=0$, 10^{-4}, 10^{-3} and 10^{-2} for chamber radius R, volume V_c and $\sigma/b=10^{-3}$. When $s\geqslant\sigma/b$, chamber expansion causes a significant increase in ΔV for a given chamber size (see Fig. 2).

assumption of a constant volume condition $(s \sim 0)$ would provide a large (up to an order of magnitude) overestimation of the stable chamber volume. For example, 10 km^3 of basaltic magma must be supplied to (and will later be erupted from) a 6.2-km radius chamber with volume 1,000 km^3 which has the ability temporarily to increase its storage capacity by as little as 1% (Fig. 3, $s = 10^{-2}$) whereas only 1 km^3 is required to trigger eruption from the same chamber if this additional space cannot be made available (Fig. 3, $s = 0$). Furthermore, caldera collapse or other tectonic effects will, by reducing the chamber volume to less than V_c, cause V_c to exceed ΔV. These effects have influenced activity in the Icelandic rift zones[32]. As V_c and particularly s, may vary over even a short span of a volcano's history, such tectonic influences will account for some of the variability between the amounts of magma erupted from the same volcano in successive eruptions.

Although basaltic magmas in plutonic conditions are generally believed to be undersaturated in volatile components, some eruptions may be related to chambers in which a free volatile phase has evolved. Because a gas is much more compressible than liquid magma, a given increase in P can be obtained by introducing a relatively small volume of magma or a relatively large volume of vapour. The critical value of ΔV is therefore virtually unaffected by the presence of a gas phase. Eruptions which are triggered by magma mixing[10] are therefore deduced to be triggered solely by the increased fluid pressure resulting from injection of (ΔV) new magma and not by any vesiculation process. Basalt–rhyolite mixed magmas have been erupted irrespective of the water content of the rhyolitic magma. Although the rhyolitic liquids involved in the Askja 1875 and Novarupta 1912 mixed pumice eruptions are believed to have been at or near saturation with H_2O contents of 3 and 5–6 wt%, respectively[33,34], much drier mixed lava flows are not uncommon[6].

The mechanisms discussed above serve as a general model on which to base concepts of stratovolcano development and seafloor spreading rates. Active periods, separated by repose periods of less than ten to hundreds of years and characteristic of a given volcano[35] are regarded as being associated with the uprise of magma at rates of several cubic metres per second over periods lasting up to several years, during which the critical amount of magma may be accumulated and eruption triggered more than once. As $\Delta V \simeq V_c$, volcano growth rates are equivalent to the upward magma flux rate beneath the volcano which is in turn governed by magma generation, accumulation and uprise dynamics. When related to an absolute time scale, the evolution and growth of stratovolcanoes and mid-ocean ridges are a potential tool with which to examine these magmatic processes[1–4,36–39].

I thank L. Wilson, H. Pinkerton and R.S.J. Sparks for comments and the NERC for a research studentship.

Received 16 October; accepted 24 December 1980.

1. Newhall, C. G. *J. Volcanol. geotherm. Res.* **6,** 61–83 (1979).
2. Katsui, Y. *et al. J. Fac. Sci., Hokkaido Univ.* Ser. IV, **16,** 533–552 (1975).
3. Rose, W. I. *et al. J. Geol.* **85,** 63–87 (1977).
4. Fairbrothers, G. E. *et al. Contr. Mineral. Petrol.* **67,** 1–9 (1978).
5. Anderson, A. T. *J. Volcanol. geotherm. Res.* **1,** 3–33 (1976).
6. Eichelberger, J. C. *Bull. geol. Soc. Am.* **86,** 1381–1391 (1975); *Nature* **275,** 21–27 (1978).
7. Sakuyama, M. *J. Volcanol. geotherm. Res.* **5,** 179–208 (1979).
8. Donaldson, C. H. & Brown, R. W. *Earth planet. Sci. Lett.* **37,** 81–89 (1977).
9. Rhodes, J. M. *et al. Tectonophysics* **55,** 35–61 (1979).
10. Sparks, R. S. J. *et al. Nature* **267,** 315–318 (1977).
11. O'Hara, M. J. *Nature* **266,** 503–507 (1977).
12. Usselman, T. M. & Hodge, D. S. *J. Volcanol. geotherm. Res.* **4,** 265–281 (1978).
13. Sparks, R. S. J. *et al. Earth planet. Sci. Lett.* **46,** 419–430 (1980).
14. Huppert, H. E. & Sparks, R. S. J. *Nature* **286,** 46–48 (1980).
15. Mogi, K. *Bull. Earthq. Res. Inst. Tokyo Univ.* **36,** 99–134 (1958).
16. Fiske, R. S. & Kinoshita, W. T. *Science* **165,** 341–349 (1969).
17. Bjornsson, A. *et al. Nature* **266,** 318–323 (1977).
18. Bjornsson, A. *et al. J. geophys. Res.* **84,** 3029–3038 (1979).
19. Wadge, G. *J. Volcanol. geotherm. Res.* **2,** 361–384 (1977).
20. Brandsdottir, B. & Einarsson, P. *J. Volcanol. geotherm. Res.* **6,** 197–212 (1979).
21. Lama, R. D. & Vutukuri, V. S. *Handbook on Mechanical Properties of Rocks* Vol. 3 (Trans. Tech. Publications, Clausthal, FRG, 1978).
22. Burnham, C. W. *Earth Miner. Sci.* **41,** 69–70 (1972).
23. Elder, J. W. *The Bowels of the Earth* (Oxford University Press, 1976); *Pure appl. Geophys.* **117,** 3–33 (1979).
24. Murase, T. & McBirney, A. R. *Bull. geol. Soc. Am.* **84,** 3563–3592 (1973).
25. Marsh, B. D. *J. Geol.* **87,** 687–713 (1979).
26. Einarsson, P. *Bull. Volcanol.* **41,** 187–195 (1978).
27. Dzurisin, D. *et al. J. Volcanol. geotherm. Res.* **7,** 241–269 (1980).
28. Swanson, D. A. *et al. U.S. geol. Soc.* Prof. Pap. 963 (1976).
29. Decker, R. W. *et al. Bull. Volcanol.* **29,** 721–732 (1966).
30. Wadge, G. *J. Volcanol. geotherm. Res.* **1,** 237–263 (1976).
31. Johnsen, G. V. *et al. J. Geophys.* **47,** 132–140 (1980).
32. Sigurdsson, H. & Sparks, R. S. J. *Nature* **274,** 126–130 (1978); *Bull. Volcanol.* **41,** 149–167 (1978).
33. Sigurdsson, H. & Sparks, R. S. J. *J. Petrol.* (in the press).
34. Clocchatti, R. & Basset, A. M. *Eos, Madr.* **59,** 225 (1978).
35. Wickman, R. E. *Ark. Miner. Geol.* **4,** 291–367 (1966).
36. Baksi, K. & Watkins, N. D. *Science* **180,** 493–496 (1973).
37. Wadge, G. *et al. Nature* **255,** 385–387 (1975).
38. Lewis, B. T. R. *Geophys. Res. Lett.* **6,** 753–756 (1979).
39. Flower, M. F. J. *Nature* **287,** 530–532 (1980).

Seismic quiescence in the Western Hellenic Arc may foreshadow large earthquakes

M. Wyss

CIRES, University of Colorado, Boulder, Colorado 80309, USA

M. Baer

Institute of Geophysics, ETH-Hoenggerberg, CH-8093 Zürich, Switzerland

South of Crete, the seismicity rate was constant in space and time between 1950 and 1978, but in the Western Hellenic Trench the rate was the same as that south of Crete until 1962, whereupon it decreased by 80%. We now suggest that the anomaly constitutes a warning of large earthquakes within the coming decade.

Studies[1,2] of the events preceding 20 earthquakes of magnitude greater than 6.0 have shown decreases of about 50% in the frequency of small earthquakes for several years before the main shocks. This was for example the case in the years preceding a Hawaiian earthquake of magnitude 7.2, where the decrease in seismicity was accompanied by two other precursor phenomena—relaxation of crustal strain and a change of seismic velocity[3]. Our hypothesis is therefore that a significant decrease of seismicity rate within a given volume is a sign that preparations have begun for one or more major earthquakes within that volume.

The region surrounding the Aegean Sea has the highest seismicity rate in western Eurasia[4], and its tectonic complexity has been widely discussed[5–11]. There is general agreement that the Hellenic trench–island arc system is a clear tectonic unit, that subduction occurs along it and that it can be regarded as a plate boundary. The direction of convergence is roughly perpendicular to the Western Hellenic Arc and approximately parallel to its eastern segment[6]. The rate of convergence may increase from west to east and is estimated to be between 2.5 and 7 cm yr^{-1} (ref. 6). Assuming that no substantial aseismic creep occurs, these consumption rates imply that enough strain energy has been stored here during loading cycles of the order of 100–200 yr to produce large earthquakes.

The time of the last very large ruptures along the Hellenic plate boundary is a matter of debate. Eight earthquakes of magnitude $\geqslant 7.75$ occurred in this area between 1805 and 1926, but most authors consider that these events occurred at depths of 100 km and locations about 100 km north of the plate boundary[12]. By this interpretation the plate boundary has not been ruptured in great events for about 200 yr, which would imply that 5–15 m of potential displacement have been stored. However, we propose the data on the destruction[18] caused by these large earthquakes to mean that these events were shallow ruptures of the plate boundary[20]. According to our interpretation, all the Hellenic plate boundary, except a 150-km segment west of Crete, was ruptured between 1805 and 1926.

0028-0836/81/090785—03$01.00

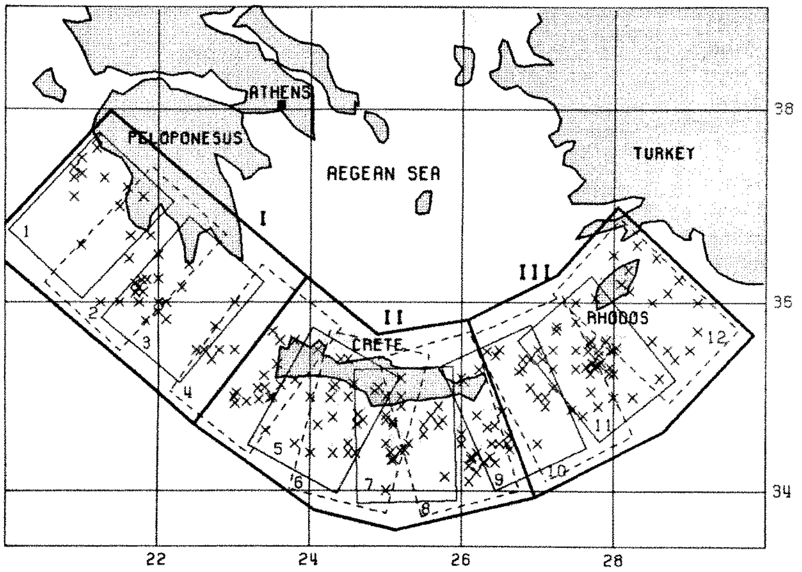


Fig. 1 Epicentre map of all earthquakes used (1950–78, 5.0 mag, $0 \leq$ depth ≤ 100 km) in the seismicity study of the Hellenic arc. Subvolumes for which seismicity rates as a function of time were determined separately are labelled 1–12. Large segments of the arc for which cumulative seismicity is shown as a function of time in Fig. 2 are labelled I, II and III.

We therefore estimate the potential earthquake slip stored at present to be $\sim$2.5–6 m (depending on the location along the arc). We conclude that, according to either interpretation, the Hellenic arc could reasonably be expected to be capable of very large earthquakes at the present time.

To determine whether any part of the arc is in a preparatory stage to rupture we studied the rate of occurrence of shallow earthquakes (0–100 km depth) as a function of space and time. A complete, pooled[13–15] data set for earthquakes of >4.9 mag was obtained for the period 1950–78. Analysis of cumulative numbers of earthquakes as a function of time showed that their overall rate of incidence in the Hellenic Arc decreased slightly after 1962. To determine the portion of the plate boundary responsible for this change we examined seismicity rates within 12 random overlapping volumes, defined in Fig. 1. Three facts became evident: (1) The seismicity rates in the neighbouring volumes 5, 6, 7 and 8 were the same to within 10%; they were approximately constant over the 28-yr period; and they were relatively high (1.5 events yr^{-1}) for this island arc. (2) Volumes 1–3 showed a common pattern of high rates ($\sim$1.5 events yr^{-1}) for the first 12 yr followed by a much lower rate (0.5 events yr^{-1}) during the remainder of the period. (3) A mixture of rates and patterns existed in volumes 9–12. Based on these observations, subvolumes of the arc with similar patterns were merged for further study of the seismicity rates.

Figure 2 shows the cumulative numbers of earthquakes as a function of time for the super-volumes designated I, II and III in Fig. 1 which are of similar dimensions (Fig. 2). The seismicity rate in these three volumes was the same up to the early 1960s, after which time volume II continued to produce earthquakes at the same rate as before, while volume I showed a highly significant decrease of seismicity to about 20% of its previous rate. We interpret this decrease of seismicity to be an anomaly precursory to one or several large earthquakes in volume I. The data of volume III are less clear, but an anomaly may have existed there since about 1970. We conclude that the greatest likelihood of very large earthquakes exists in the western part of the Hellenic Trench.

The largest length of the arc segment which may be ready for rupturing is approximately 400 km (volume I and immediately adjoining volumes containing high seismicity). Based on the destruction caused by the great earthquakes which occurred in the Western Hellenic Trench during the nineteenth and early twentieth centuries[18], we estimate the most likely length of future ruptures to be about 100 km. Using the hypothesis that the seismicity anomaly in the Western Hellenic Trench represents a precursor, and assuming that the precursor time is correlated with anomaly dimension, we expect that the sequence of large earthquakes will start between 1980 and 1991. Using seismicity data alone, we cannot estimate the time of occurrence more accurately than this because of the scatter in the data defining the anomaly dimension versus precursor time relation[1].

We expect an earthquake with the following source parameters to rupture the Hellenic arc plate boundary: location, near 23.5°E; size, lengths = 100 ± 50 km, magnitude 7.75 ± 0.5; occurrence, $1980 \leq T_0 \leq 1991$.

Although precursor seismicity patterns cannot be identified with certainty along the other segments of the Hellenic Trench, it seems reasonable to expect large earthquakes there also. However, this prediction should not be made without emphasizing the rudimentary nature of our understanding of the processes which lead to earthquakes. Even though the number

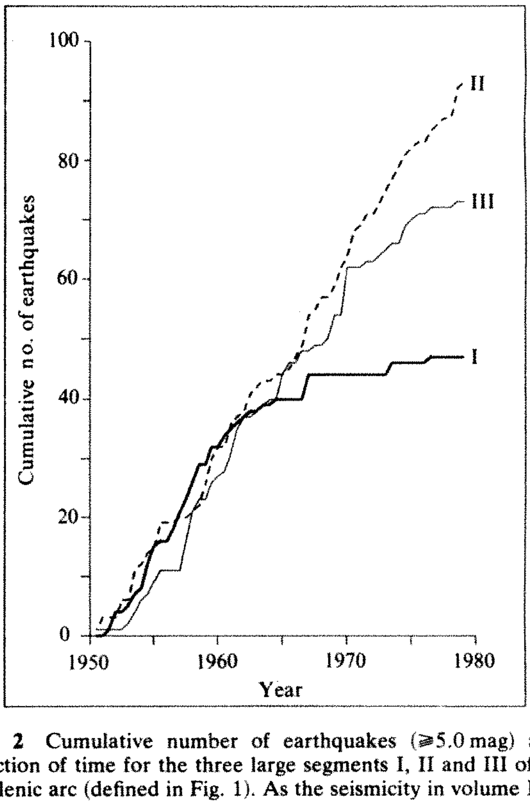


Fig. 2 Cumulative number of earthquakes (≥ 5.0 mag) as a function of time for the three large segments I, II and III of the Hellenic arc (defined in Fig. 1). As the seismicity in volume III is dominated by swarms (steps in the curve), we can reach no conclusion for this region. Note that the seismicity rates in I and II are closely similar up to 1962, after which time the rate in volume I is reduced to 20% of this rate while the rate in II continues at the pre-1962 level. The establishment of the World Wide Standard Seismograph network did not affect seismicity resolution in the area because the location capability is mostly based on European stations.

of well documented seismicity quiescence precursors is increasing, we do not know whether it is possible for seismicity quiescence anomalies to be terminated without a main shock.

Nevertheless, we believe that precautionary measures are warranted, although we are aware of the possible negative effects of broadcasting an earthquake prediction[19]. Ohtake *et al.*[16], about 2 yr before the 7.8 mag Oaxaca earthquake in 1978, showed that quiescence had begun in the Oaxaca gap in mid-1973, but their report produced no constructive action and was, in fact, ultimately detrimental to the area concerned. We hope that previous mistakes will not be repeated in the case of the Hellenic Trench and that scientific surveillance of the area will be started.

This research was supported by funds of the Institute for Geophysics, ETH-Zürich, and by US Geological Survey contract no. 14-08-001-18386.

Received 14 October; accepted 2 December 1980.

1. Wyss, M. & Habermann, R. E. *Pure appl. Geophys.* **117,** 1195 (1979).
2. Habermann, R. E. *3rd Maurice Ewing Symp.* Abstr. (1980).
3. Wyss, M., Klein, F. W. & Johnston, A. C. *Nature* **289,** 231–234 (1981).
4. Barazangi, M. & Dorman, J. *Bull. seism. Soc. Am.* **59,** 369 (1969).
5. McKenzie, D. P. *Nature* **226,** 239 (1970); *Geophys. J. R. astr. Soc.* **55,** 217 (1978).
6. Le Pinchon, X. & Angelier, J. *Tectonophysics* **60,** 1 (1979).
7. Papazachos, B. C. & Delibasis, N. D. *Tectonophysics* **7,** 321 (1969).
8. Papazachos, B. C. & Comninakis, P. E. *J. geophys Res.* **76,** 8517 (1971).
9. Papazachos, B. C. *Geophys. J. R. astr. Soc.* **33,** 421 (1973).
10. Comninakis, P. E. & Papazachos, B. C. *Bull. geol. Soc. Am.* **83,** 1093 (1972).
11. Berckhemer, H. in *Int. Symp. Structural History of the Mediterranean Basins* (eds Eeju-Duval, B. & Montadert, L.) 303 (Editions Technip, Paris, 1977).
12. Galanopoulos, A. *Ann. Geofis.* **20,** 109 (1967).
13. Comninakis P. E. & Papazachos, B. C. *A Catalogue of Earthquakes in the Mediterranean and Surrounding Area for the Period 1901–1975* (Geophysical Laboratory, University of Thessaloniki, 1978).
14. Drakopoulos, J. *et al. Bull. natn. Obs. Athens*, Seism. Inst. (1978).
15. *World Hypocenter Data File* (NOAA, Washington DC, 1979).
16. Ohtake, M., Matumoto, T. & Latham, G. V. *Pure appl. Geophys.* **115,** 375 (1977).
17. Galanopoulos, A. *Ann. Geofis.* **20,** 109 (1967).
18. Sieberg, A. *Denkschr. med.-naturw. Ges. Jena* **18,** 2nd edn.
19. Garza, T. & Lomnitz, C. *Pure appl. Geophys.* **117,** 1187 (1979).
20. Wyss, M. & Baer, M. *3rd Maurice Ewing Symp.* Abstr. (1980).

Miocene spreading centre south of Isla Guadalupe

Rodey Batiza

Department of Earth and Planetary Sciences and
McDonnell Center for Space Sciences, Washington University,
St Louis, Missouri 63130, USA

Clement G. Chase

Department of Geology and Geophysics, University of Minnesota,
Minneapolis, Minneasota 55455, USA

The Juan de Fuca, Rivera, Cocos and Nazca plates in the eastern Pacific are small fragments remaining from the fragmentation and partial subduction of the long, narrow Farallon Plate which began as early as 55 Myr ago[1]. This fragmentation apparently resulted from interaction of the very irregular Pacific–Farallon ridge with the uneven coastlines of the American Plate[1–4] and involved the formation of ridge–trench transforms[5], creation of new spreading centres[6], ridge jumps[1,7,8], deactivation of subduction zones[2] and other phenomena[1]. We present here magnetic evidence that a fifth small fragment of the Farallon plate is preserved in the eastern Pacific off the coast of Baja, California. Its preservation is the result of an eastward jump of the Pacific–Farallon ridge ~12.5 Myr ago which was preceded by asymmetric spreading and perhaps pivotal subduction[1]. Menard[1] first suggested this idea and showed that anomaly 5B is symmetric and parallel to a linear trough trending southwards from Guadalupe Island. He named the small fragment of the Farallon plate preserved between this trough and the margin of Baja California to the east, the Guadalupe plate[1]. We show that both anomalies 5B and 5C symmetrically flank the Guadalupe trough. Our results and conclusions are not based on new magnetic data, but rather on reinterpretation of aeromagnetic profiles published by Taylor *et al.*[9].

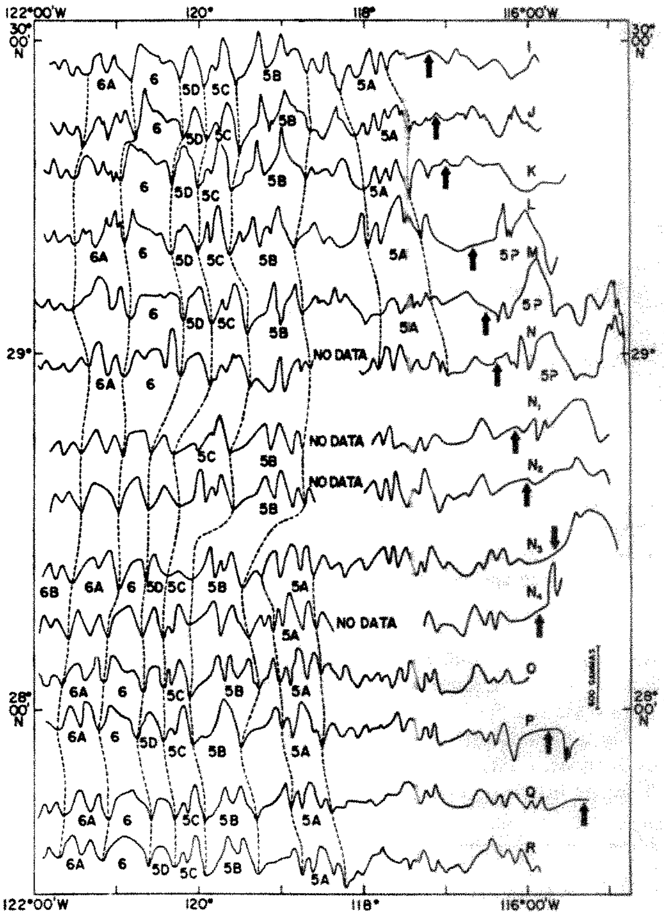


Fig. 1 Original data and interpretation of Taylor *et al.*[9] reproduced exactly. Arrows indicate the location of the 1,000-fathom bathymetric contour.

Figure 1 shows the data of Taylor *et al.*[9]. This area of the eastern Pacific was first studied by Krause[10] who recognized lineated magnetic anomalies east of Isla Guadalupe. Later studies[2,11] confirmed their presence, but as Atwater[2] states "near Guadalupe Island, (the magnetics) lack distinctive character and thus remain unidentified except that they probably belong to the anomaly 5B-5A sequence." Menard[1] was apparently confident that anomaly 5B could be identified there but his interpretation has remained controversial. In our interpretation of the magnetic anomalies of the ares we rely on closely spaced east–west aero magnetic anomaly profiles[9] which are shown together with their original interpretation in Fig. 1. These data allow the anomalies in the unevenly preserved anomaly 5 series to be identified more confidently than before.

Taylor *et al.*[9] emphasized the variable seafloor spreading rates in the vicinity of Isla Guadalupe. In their interpretation, spreading rates decrease monotonically from the latitude of Isla Guadalupe northwards and increase southwards although remaining at uniformly lower rates than north of the island. Their suggestion that a fracture zone is present at about the latitude of Isla Guadalupe has been confirmed by seismic refraction work[12]. However, Taylor *et al.*[9] did not attempt a detailed reconstruction of the tectonic history of the area.

We have renumbered the magnetic anomalies of Taylor *et al.*[9] in accordance with a more recent magnetic time scale and anomaly nomenclature[13]. Figure 2 shows our new anomaly identifications which differ from those of Taylor *et al.*[9] (Fig. 1) by the addition of anomaly 5E and the shift of the rest of the anomaly 5 series. We have also reinterpreted some of the profiles south of Isla Guadalupe and have identified reversed anomalies 5B and 5C to the east of Guadalupe (Fig. 3). Finally, we have plotted the magnetic anomalies of Fig. 1 onto Taylor *et al.*'s[9] map of the area to show the geographical distribution of

0028-0836/81/090787—03$01.00

Table 1 Average spreading rates

	6A–5E (West)		5D–5C (West)		5C–5B (West)		5B–12.56–5B (Fossil ridge)		5B–5C (East)	
	Profile	Rate (mm yr^{-1})	Profile	Rate (mm yr^{-1})	Profile	Rate (mm yr^{-1})	Profile	Rate (mm yr^{-1})	Profile	Rate (mm yr^{-1})
North segment	CC	30								
	FF	31	B	50						
	B	32								
	D	32	D	50	D	50?				
	G	33								
Middle segment	K	35	K	48?	J	46				
	M	34	L	46?	K	46				
			M	48?						
South segment	N_2	33	N_2	40	N_3	43	N	34	N_2	46
	O	30	N_3	40	O	41	N_2	34	O	44
	P	30	O	40	Q	39	N_3	36	Q	45
	Q	30	R	41	R	41	N_4	35?		
							Q	33		
							Q	32		

Profile letters correspond to those of ref. 9, some of which are shown in Fig. 1. North, middle and south segments are shown in Fig. 4. Spreading rates were estimated by comparing data profiles with synthetic profiles of known spreading rate, therefore estimates of precision of these values are subjective, and are no better than ±1 mm yr^{-1}.

the magnetic anomalies (Fig. 4). Table 1 shows average spreading rates for various time intervals computed using Blakely's time scale for the Miocene[13].

Figure 4 indicates important aspects of the tectonic evolution of the area. First, we have detected a major offset in the anomaly 5 series north of Guadalupe (Fig. 4) which corresponds to a series of east–west ridges. The magnitudes of the offsets of anomalies 5E, 5D, and 5C increases monotonically as they do for the fracture zone at the latitude of Isla Guadalupe. Several smaller anomaly offsets may also exist in the region north of the island, but more closely spaced magnetic data will be needed to confirm their presence. Figure 4 represents the simplest possible anomaly pattern which can be derived from the data.

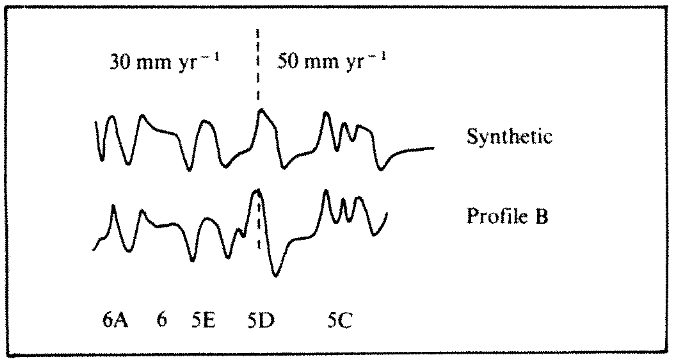


Fig. 2 Comparison of synthetic anomaly profile with data of profile B. Our new identifications are shown. Note variable spreading rate of synthetic profile.

The tectonic history of the region inferred from this pattern of anomalies is: (1) During anomaly 6A and 6 time (20–22.3 Myr BP) two segments of the Pacific–Farallon system offset by a transform fault at about the latitude of Isla Guadalupe were spreading at maximum rates of ~35 mm yr^{-1} (Table 1). The rates for the northern segment increase southwards producing fanning anomalies and an arcuate fracture zone. For this segment, rates increase from ~30 to 33 mm yr^{-1} in the south. Anomaly 6A shows little or no offset, but anomaly 6 is offset across the fracture zone; thus, either spreading was asymmetric or the northern and southern segments belonged to different plates. We prefer the first explanation.

(2) Between anomaly 5E and 5B (12.5–20 Myr BP) reorganization of spreading geometry resulted in formation of a third ridge segment north of Isla Guadalupe. This may have been caused by interaction of this portion of the Farallon plate with the continental margin of Baja California. Anomaly offsets across this transform fault–fracture zone increased monotonically with time and spreading rates were variable (Table 1). Spreading rates south of Isla Guadalupe increased to about 40 mm yr^{-1} as shown previously[9], but then slowed again to about 35 mm yr^{-1}.

(3) Just after anomaly 5B time (12.5 Myr BP) the southernmost segment of the Pacific–Farallon system in this area jumped to the east. This is inferred from the presence of anomalies 5B and 5C east of the Guadalupe trough, which we suggest is an extinct spreading centre. This series of discontinuous linear troughs (Fig. 4) extends from the southern tip of Isla Guadalupe to the Shirley Trough at about 27°N (J. Mammerickx, personal communication). Table 1 shows that during anomaly 5C and 5B time, spreading at the Guadalupe Trough was asymmetric: about 41 mm yr^{-1} on the west flank and 45 mm yr^{-1} on the east. This is why we favour asymmetric spreading over plate

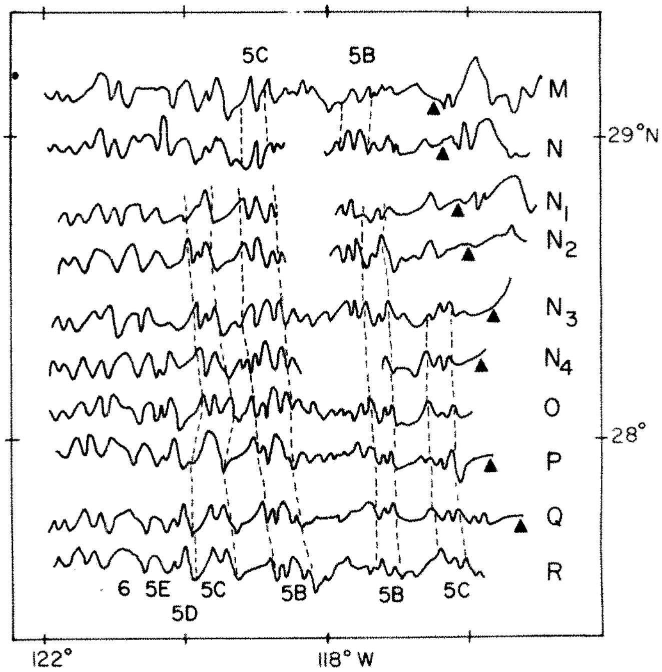


Fig. 3 Dashed lines show our new magnetic anomaly identifications and correlations for the southern part of the area which includes the Guadalupe Trough. ▲, Location of the 1,000-fathom bathymetric contour.

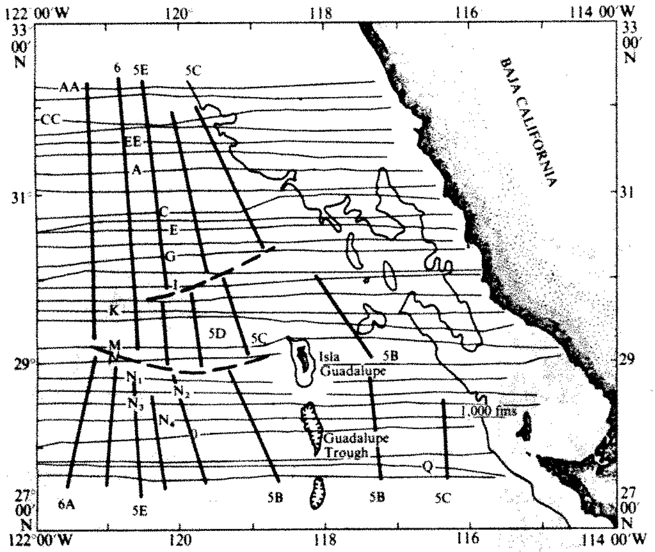


Fig. 4 Our new magnetic anomaly interpretations and the Guadalupe Trough[14] added to one of Taylor *et al.*'s maps.

fragmentation as an explanation for the smoothly increasing magnetic anomaly offsets on both transform fracture zone systems in the area.

To test our new magnetic anomaly interpretation east of Isla Guadalupe as well as the hypothesis of interrupted spreading (ridge jump), synthetic mangetic profiles for continuous versus interrupted spreading are compared with data of Taylor *et al.*[9] in Fig. 5. While neither synthetic is a perfect match to the data, the interrupted spreading model is certainly a better fit.

It is not known whether the ridge segments north of Guadalupe also jumped to the east at the time when the Guadalupe Trough became inactive because linear magnetic anomalies farther east have not been recognized. If Taylor *et al.*[9] are correct in their identification of anomaly 5 in profiles L, M and N then these segments probably continued to spread continuously (at least until 9 Myr BP). Alternatively, Taylor *et al.*'s[9] anomaly 5 could be 5A, or a buried basement ridge, or perhaps even another fossil ridge crest. In the former case, Isla Guadalupe occupies the intersection of a fracture zone and a fossil ridge, and is not simply built along a fossil ridge as previously proposed[14]. More detailed magnetic data will be required to determine the precise location of this fraction zone and thus the actual tectonic environment in which the Guadalupe volcanoes originated and evolved.

The spreading rate data of Table 1 give an insight into the possible relationships between the observed systematic changes of spreading rate before the ridge jump and the underlying causes of such jumps. As shown above, the ridge jump was probably preceeded by at least 10 Myr of asymmetric spreading in the entire region. Spreading rates increased from 30–33 to 40–50 mm yr^{-1} between anomaly 5E and 5D time (18–19 Myr) and remained high but variable until just before the ridge jump when spreading rates again slowed to ~34 mm yr^{-1}. While the underlying causes of such jumps are not yet well understood, clues to the mechanism for ridge jumps may be provided by detailed study of the spreading history of a ridge segment before episodic migration. In the case of the Guadalupe trough, the slower spreading, its asymmetry and the interaction between the farallon plate and the continental margin of Mexico may all have contributed to the ridge jump. Alternatively, the spreading history of the Guadalupe trough may reflect the action of different causes of ridge migration, for example, relative displacement between elements of mantle convention and the overlying lithosphere.

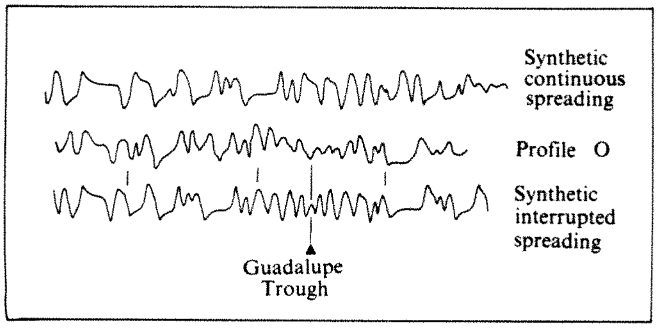


Fig. 5 Comparison of profile O data to both continuous and interrupted spreading synthetic anomaly profiles. The interrupted spreading synthetic is a composite of four profiles generated at various spreading rates (segments separated by short vertical lines above the profile). The spreading rates for these segments are: 30 mm yr^{-1} (west), 40 mm yr^{-1}, 33 mm yr^{-1} and 44 mm yr^{-1} (east). Although neither synthetic is a perfect match to the data, interrupted spreading is a much better fit.

We thank Patrick Taylor, Tanya Atwater and Richard Hey for discussions, and J. Mammerickx for an unpublished chart. This work was partly supported by NSF grant OCE-7727001 and an ONR grant to Rodey Batiza and an NSF grant to Chase.

Received 7 October; accepted 11 December 1980.

1. Menard, H. W. *J. Geol.* **86,** 99–110 (1978).
2. Atwater, T. *Bull. geol. Soc. Am.* **86,** 3513–3536 (1970).
3. Herron, E. M. *Bull. geol. Soc. Am.* **83,** 1671–1692 (1972).
4. Handschumacher, D. W. *Geophys. Mongr.* **19,** 177–202 (1976).
5. Molnar, P. *Bull. geol. Soc. Am.* **84,** 1651–1658 (1973).
6. Herron, E. M. & Heirtzler, J. R. *Science* **158,** 775–780 (1967).
7. Anderson, R. N. & Davis, E. E. *Nature* **241,** 191–193 (1973).
8. Sclater, J. G., Anderson, R. N. & Bell, M. L. *J. geophys. Res.* **76,** 7888–7915 (1971).
9. Taylor, P. T., Brennan, J. A. & O'Neill, N. J. *Nature* **229,** 396–399 (1971).
10. Krause, D. C. *Bull. geol. Soc. Am.* **76,** 617–650 (1965).
11. Chase, C. G., Menard, H. W., Larson, R. L., Sharman, G. F. & Smith, S. M. *Bull. geol. Soc. Am.* **81,** 491–498 (1970).
12. Spudich, P. & Orcutt, J. *J. geophys. Res.* **85,** 1409–1433 (1980).
13. Blakely, R. J. *J. geophys. Res.* **79,** 2974–2985 (1974).
14. Batiza, R. *Geology* **5,** 760–764 (1977).

Vertical distribution of mercury in sediments from the East Pacific Rise

Malcolm E. Cox & Gary M. McMurtry

Hawaii Institute of Geophysics, University of Hawaii, Honolulu, Hawaii 96822, USA

Concentrations of mercury are often high in waters and sediments associated with volcanic hydrothermal systems, in both marine and terrestrial environments. Whereas the concentration of mercury in open ocean waters is reported to be in the range 0.002–0.05 µg l^{-1} (refs 1–4), significantly higher values of up to 1.4 µg l^{-1} are reported in waters near, or having circulated through, areas of submarine hydrothermal activity[2,5]. Mercury concentrations of up to 400 p.p.b. (parts per 10^9) on a carbonate-free basis (CFB) have been measured in surface sediments on the East Pacific Rise crest, but concentrations decrease to <20 p.p.b. (CFB) on the peripheral seafloor[6]. Similar high concentrations from 160 to 890 p.p.b. (CFB) have been recorded in near-surface sediments from the median valley of the Mid-Atlantic Ridge[7]. We report here on the vertical variations of mercury in sediments from the East Pacific Rise, and suggest that increased concentrations are correlated with periods of increased submarine volcanic activity.

Detailed down-core mercury analyses were made at 10-cm intervals on three piston cores collected by the RV Kana Keoki in the vicinity of the East Pacific Rise. Two of the cores, KK71-PC85 and KK71-PC87, located within 40 km of each other, are from within 50 km of the palaeomagnetically defined spreading axis of the East Pacific Rise at 6° S latitude; the third, KK72-PC8, is from the adjacent Bauer Basin (Fig. 1). These analyses were conducted to observe what vertical variations may occur in mercury concentration in marine sediments, knowing

0028-0836/81/090789—04$01.00

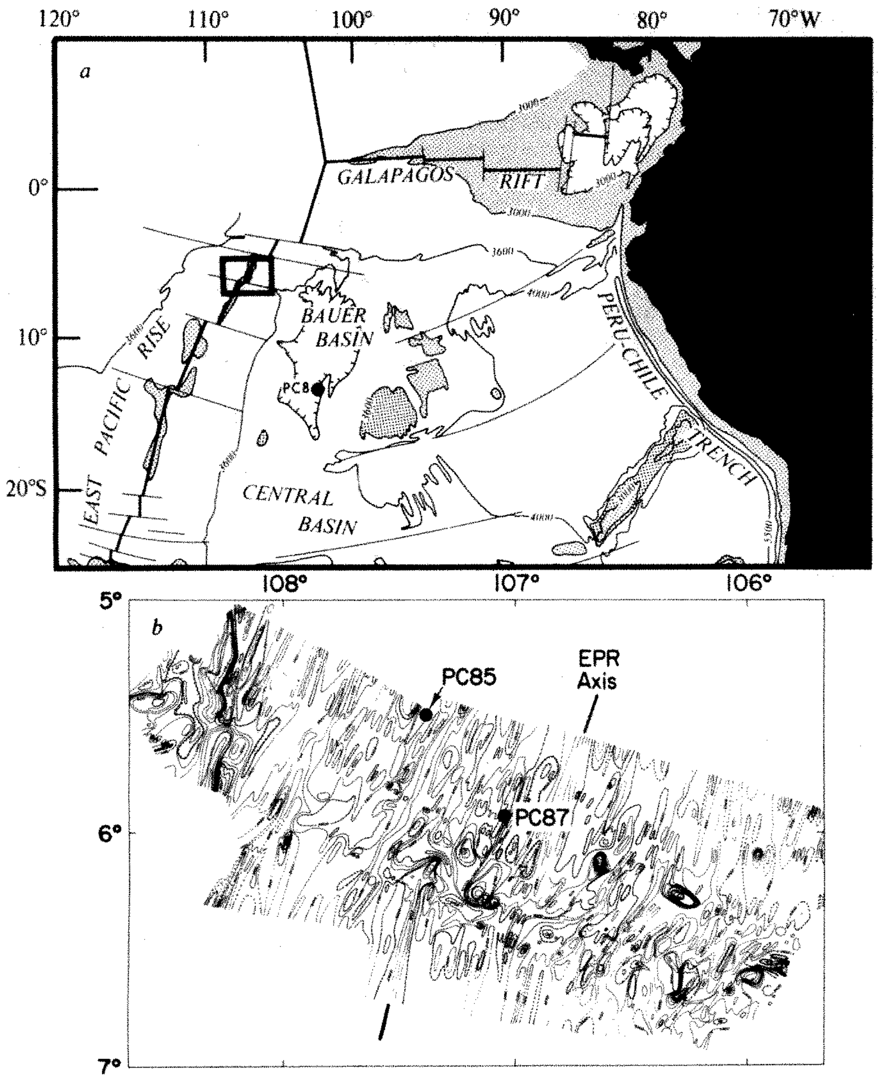


Fig. 1 *a*, General bathymetric map of the Nazca plate, south-east Pacific, showing major physiographic features and core locations. Box indicates detail map (*b*). Bathymetry after ref. 29. *b*, Detailed site survey area of the East Pacific Rise at 6° S latitude. Bathymetry after ref. 30.

that (1) both terrestrial and submarine volcanism are episodic in nature; (2) temporal variations in bottom water mercury anomalies over spreading centres have been observed[8]; and (3) most published sediment mercury analyses are from the surface or near-surface.

Samples were immediately taken from the split core sections and were individually sealed in plastic vials before analysis on a Jerome Instrument Corporation model 301 gold film mercury detector. Weighed, air-dried samples were heated to >800 °C, and the mercury vapour released was transported by Hg-free carrier gas through various filters to remove water and acid vapours before adsorption onto the gold film. Changes in the resistance of the gold film due to mercury adsorption were monitored and calibrated by injections of standard amounts of

Table 1 Down-core chemical composition

Core no.	General lithological description	Sedimentation rate* (cm kyr^{-1})	$CaCO_3$† (wt %)	SiO_2 (wt %)	Fe_2O_3 (wt %)	MnO (wt %)	Al_2O_3 (wt %)	Cu (p.p.m.)	Zn (p.p.m.)	Hg‡ (p.p.b.)
KK71-PC85	Radiolarian-rich, nannofossil foram ooze	2.4 ± 0.3	82.9 ± 3.3	12.7 ± 2.9	2.00 ± 0.62	0.17 ± 0.05	0.24 ± 0.13	268 ± 64	36 ± 16	58 ± 44
			n§ = 12							*n* = 48
KK71-PC87	Clay, radiolarian and diatom bearing, foram-rich, nannofossil ooze. Minor ash bands throughout core	1.4 ± 0.1	81.3 ± 4.1	10.7 ± 1.7	2.85 ± 1.23	0.43 ± 0.35	0.47 ± 0.61	268 ± 42	34 ± 14	29 ± 29
			n = 18							*n* = 80
KK72-PC8	Nannofossil bearing, ash-rich clay	0.41 ± 0.06	1.39 ± 0.14	30.9 ± 4.1	27.7 ± 3.8	7.25 ± 1.24	8.28 ± 0.84	1061 ± 97	113 ± 24	0.7 ± 1.8
			n = 22							*n* = 42

* Determined by excess ^{230}Th method using γ-ray spectrometry[17].

† $CaCO_3$ determined by gas volume method; all others by X-ray fluorescence spectrometry. Precision for $CaCO_3$ ranges from 1 to 6%; all others within 6% except Cu (15%) and Zn (50%) (ref. 27).

‡ Hg analyses performed by gold-film mercury analyser[9] which compares with analyses by atomic absorption[28]. Precision within 5%. Accuracy within 10%.

§ No. of samples.

mercury saturated air[9]. The accuracy of the procedures was monitored with a soil standard (USGS geochemical exploration sample no. 5) of known mercury concentration. General lithogical descriptions, sedimentation rates and mean down-core chemical compositions based on samples taken at 50-cm intervals are presented along with the mean down-core mercury results in Table 1. Figure 2 shows plots of mercury concentration versus depth for the East Pacific Rise sediment cores.

The extreme vertical variations in mercury concentration in the East Pacific Rise sediments are significant, and may demonstrate episodic submarine volcanic activity. In agreement with previous work[7], the variations in mercury show little correlation to variations in other sediment components. By using an approximate sedimentation rate of 2 cm kyr^{-1}, and assuming that each major mercury peak represents an episode of significant submarine volcanic activity and that mercury concentrations are not appreciably affected by variable bottom-water movements, we can calculate mean recurrence intervals of the order of 25,000–35,000 yr for this region of the East Pacific Rise. These intervals agree within factors of two or three with periods separating major cycles of spreading activity (~10,000 yr) suggested from geological evidence for the FAMOUS area, the Galapagos Rift[10] and the East Pacific Rise at 21° N latitude[11]. However, this correlation further assumes that rise crest volcanic activity is confined to the spreading axis, and does not consider such activity or hydrothermal emission from nearby faults and seamounts (Fig. 1) which may complicate the observed vertical distribution of mercury. The lack of correspondence between major mercury peaks in the two cores and the generally higher concentration in core KK71-PC85, which is more distant from the spreading axis, suggest that the volcanic activity responsible for the mercury concentrations may be localized. More detailed sampling will be needed to establish the spatial and temporal relationship between volcanic activity and sediment mercury concentration. These measured variations in mercury concentration are also important in illustrating that surface values alone are not an adequate representation of the mercury concentration at a given location.

The absence of mercury in the Bauer Basin core agrees with previous surface work in this region[6]. Of the 42 analyses over the topmost 3.5 m of core KK72-PC8, only seven were above the 1 p.p.b. detection limit (Table 1). The basin is characterized by sediments rich in metals[12–14], and anomalously high heat flow has been measured there[15]. Various origins for the metals in these sediments have been proposed, including transport of hydrothermally-derived material into the basin from the adjacent East Pacific Rise[16], and hydrothermal activity within the Bauer Basin[14,15]. If a hydrothermal origin is feasible, then the fixation process for mercury in these sediments must be different from that for East Pacific Rise sediments. We propose that this difference results from a higher oxidation potential caused by the relatively lower sedimentation rates in the basin (Table 1 and ref. 17). About 80% of the inorganic mercury in oxic seawater will exist as the highly soluble HgCl$_4^{2-}$ form[18], whereas aqueous mercury will exist as the less soluble Hg^0 or HgS$_2^{2-}$ forms in less oxic waters[19,20]. This oxidation–reduction behaviour may also apply to organically complexed forms of mercury. Approximately 30% of the total aqueous mercury in seawater has been shown to associate with organic matter[21], and it is reasonable to assume that such associations will also occur in pelagic sediments. Preservation of organic matter in pelagic sediments is a function of the rate of oxidation, which is inversely related to the rate of sedimentation[22]. Similar oxidation–reduction behaviour has been proposed to explain the relative distributions of chromium[14] and of uranium[23] in the Bauer Basin and East Pacific Rise sediments.

We plan further work on submarine mercury distributions, and are currently investigating sediments from back-arc spreading centres and marine manganese deposits. On the Fiji Plateau, where mercury concentrations in surface sediments are

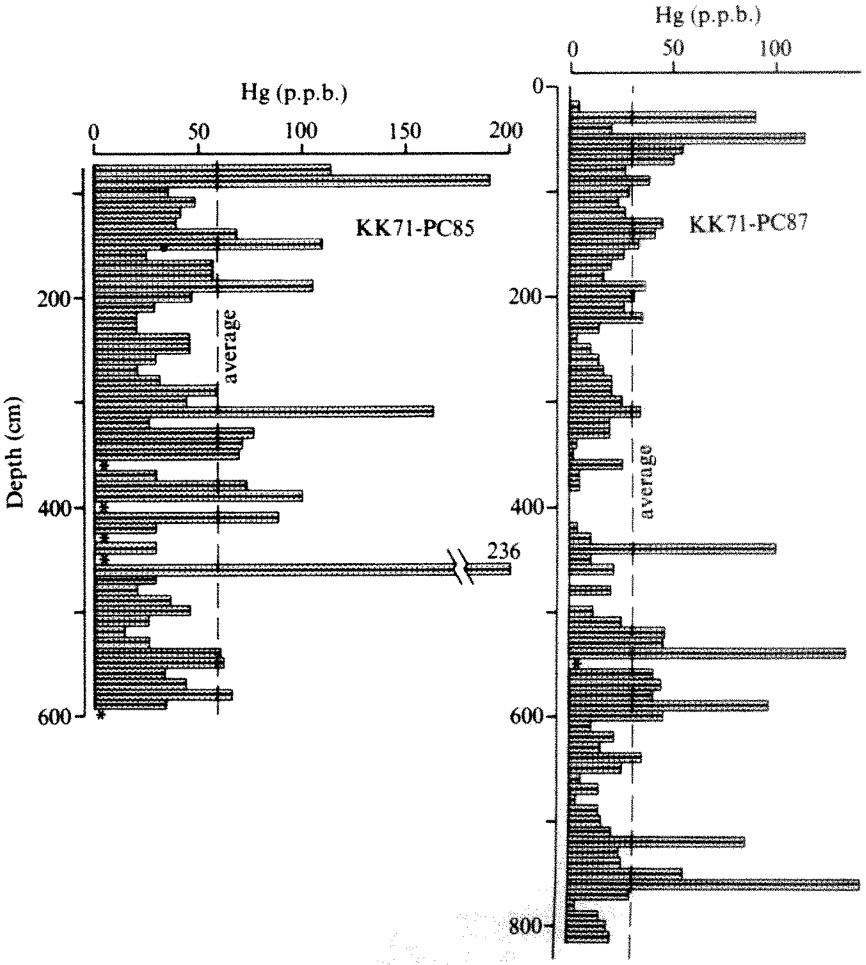


Fig. 2 Mercury concentration plotted against depth for the East Pacific Rise sediment cores. Because of coring disturbances, the topmost 70 cm of core KK71-PC85 was not analysed. Bar width indicates vertical uncertainty (±5 cm) due to sampling interval. The vertical width of each sample for mercury analysis was within ~1 cm. * Not determined.

generally low (<10–70 p.p.b.), the highest values correlate with increased heat flow[24]. Concentrations of mercury are extremely variable within marine manganese nodules (<1–810 p.p.b.), but are highest (>58 p.p.b.) in nodules from sea floor areas characterized by volcanic and hydrothermal features[25]. Similarly, mercury concentrations from several hundred to >2,000 p.p.b. were measured in ferromanganese crusts from active spreading centres[26].

This work was supported in part by the NSF, International Decade of Ocean Exploration. We thank Keith Chave and Sanford Siegel for their comments and Christine Mato for assistance in core sampling. Hawaii Institute of Geophysics contribution no. 1112.

Received 25 August; accepted 5 December 1980.

1. Leatherland, T. M., Burton, J. D., McCartney, M. J. & Culkin, F. *Nature* **232,** 112 (1971).
2. Carr, R. A., Jones, M. M. & Russ, E. R. *Nature* **251,** 489–490 (1974).
3. Fitzgerald, R. A., Gordon, D. C. Jr. & Cranston, R. E. *Deep-Sea Res.* **21,** 139–144 (1974).
4. Matsunaga, K., Nishimura, M. & Konishi, S. *Nature* **258,** 244 (1975).
5. Aidin'yan, N. Kh. *Petrog. miner. Geokhim.* **70,** 9–14 (1962).
6. Boström, K. & Fisher, D. E. *Geochim. cosmochim. Acta* **33,** 743–745 (1969).
7. Cronan, D. S. *Can. J. Earth Sci.* **9,** 319–323 (1972).
8. Carr, R. A., Jones, M. M., Warner, T. B., Cheek, C. H. & Russ, E. R. *Nature* **258,** 588–589 (1975).
9. McNerney, J. J., Buseck, P. R. & Hanson, R. C. *Science* **178,** 611–612 (1972).
10. van Andel, Tj. H. & Ballard, R. D. *J. geophys. Res.* **84,** 5390–5406 (1979).
11. Spiess, F. N. *et al. Science* **207,** 1421–1433 (1980).
12. Boström, K. & Peterson, M. N. A. *Mar. Geol.* **7,** 427–447 (1969).
13. Sayles, F. L. & Bischoff, J. L. *Earth Planet. Sci. Lett.* **19,** 330–336 (1973).
14. McMurtry, G. M. & Burnett, W. C. *Nature* **254,** 42–44 (1975).
15. Anderson, R. N. & Halunen, A. J. Jr. *Nature* **251,** 473–475 (1974).
16. Heath, G. R. & Dymond, J. *Bull. geol. Soc. Am.* **88,** 723–733 (1977).
17. McMurtry, G. M., Veeh, H. H. & Moser, C. *Geol. Soc. Am. Mem.* (in the press).
18. Anfalt, T., Dyrssen, D., Ivanova, E. & Jagner, D. *Svensk kem. Tidsskr.* **80,** 340–342 (1968).
19. Hem, J. D. *U.S. Geol. Sur. Prof. Pap.* **713,** 19–24 (1970).
20. Siegel, B. Z. & Siegel, S. M. in *The Biogeochemistry of Mercury in the Environment*, 131–159 (Elsevier, Amsterdam, 1979).
21. Windom, H. L. & Taylor, F. E. *Deep-Sea Res.* **26A,** 283–292 (1979).
22. Lynn, D. C. & Bonatti, E. *Mar. Geol.* **3,** 457–474 (1965).
23. Veeh, H. H. *Geol. Soc. Am. Mem.* (in the press).
24. Cox, M. E. *S. Pacif. Mar. Geol. Not.* **1,** no. 10, 111–122 (CCOP–SOPAC, Suva, Fiji, 1980).
25. Harriss, R. C. *Nature* **219,** 54–55 (1968).
26. Toth, J. R. *Bull. geol. Soc. Am.* **91,** 44–54 (1980).
27. McMurtry, G. M. thesis, Univ. Hawaii (1979).
28. Ohkawa, T., Uenoyama, H. & Kondo, M. *Eisei Kagaku* **22,** 11–19 (1976).
29. Mammerickx, J., Anderson, R. N., Menard, H. W. & Smith, S. M. *Bull. geol. Soc. Am.* **86,** 111–118 (1975).
30. Rea, D. K. *J. geophys. Res.* **81,** 1495–1504 (1976).

Precambrian microfossils from the Sinian of South China

Zhang Zhongying*

Department of Geology, Nanjing University, Nanjing, China

The discovery of previously unknown microfossils from the Sinian of western Hubei Province, South China (Fig. 1) is reported here. The Sinian there, up to 1,000 m thick, is naturally divisible into four formations (Fig. 2) from the base upwards: (1) the Liantuo Formation comprising feldspar–quartz–sandstones with basal conglomerates; (2) the Nantuo Formation, which is a tillite; (3) the Doushantuo Formation, which has microdolomites intercalated with black shales and cherty beds with some siliceous or phosphatic nodules; and (4) the Dengying Formation, which is mainly dolomites. The rock sample in which the microfossils were found was collected from the lower part of the Doushantuo Formation near the highway side ~500 m east of Liantuo, Yichang County (Fig. 1). The Doushantuo Formation has been dated as 693±66 Myr (Rb–Sr isochron age)[1]. Pseudomorphs after gypsum sometimes occur in the lower part. The microfossils are preserved in siliceous nodules, particularly in the cherty portions. The dark colour of the nodules is due to the high organic content.

These new Doushantuo microfossils were discovered in petrographical thin sections and include at least four types—spheroids, colonies composed of unicells, nonseptate filaments and septate filaments. In some, fine structural details are well preserved, while others are modified or destroyed by diagenetic and subsequent processes. Figure 3 shows only a few of the new Doushantuo assemblage and a more detailed systematic account will be published elsewhere.

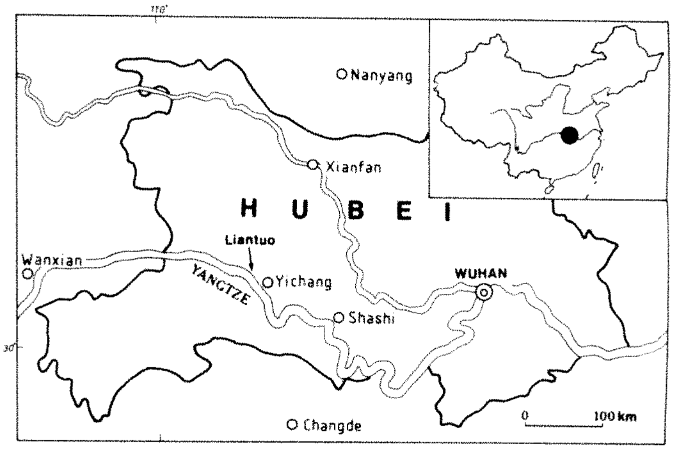


Fig. 1 Locality map.

Spheroidal members (Fig. 3*b–d*) of the assemblage range from ~8 to 18 μm in diameter, and exhibit granular or psilate surface textures. Although mainly solitary, and not associated with other organic matter, they are occasionally paired, or grouped in clusters. Some seem to contain inner dark spots ~2 μm across (Fig. 3*c*). In many colonies (Fig. 3*h*) where single spheroids 2–9 μm in diameter occur associated spheroids have also been found. Colonies may be up to >100 μm across, comprising a few to several tens of cells which set in a matrix of apparently mucilaginous material. Some of these colonies seem to have been disintegrated into smaller ones or into separate cells. These spheroids and colonies composed of unicells might represent chroococcacean cyanophytes, although some have uncertain affinities.

Both non-septate and septate filaments occur in the Doushantuo Formation, and some probable isolated trichomes and slime trails left by gliding filamentous algae or bacteria can be found. Organic walled non-septate filaments (Fig. 3 *e–g*) are generally straight or slightly sinuous, unbranched, 2–9 μm in diameter. Their maximum overall length is 56 μm. Occasionally loosely spirally twisted filaments have been found in petrographical thin sections. Variations in preservation may give these filaments a superficially septate appearance (Fig. 3*g*). Most of the non-septate filaments are interpreted to be discarded sheaths that originally enclosed cellular blue-green algal

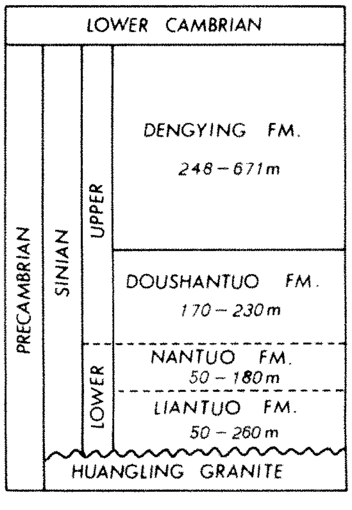


Fig. 2 Generalized stratigraphy of the Sinian in western Hubei Province, China.

---

* Present address: Department of Geology, Royal School of Mines, Imperial College, London SW7 2BP, UK.

0028-0836/81/090792—02$01.00

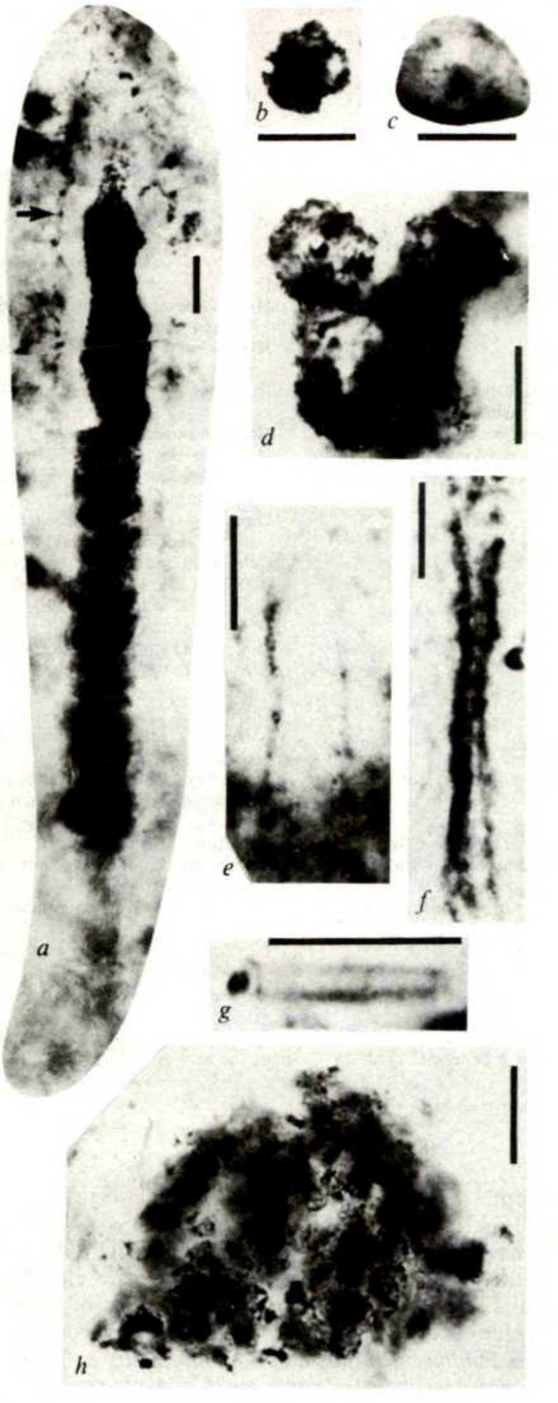


Fig. 3 Organically preserved microfossils from the Doushantuo Formation (late Sinian) of western Hubei Province, China. All micrographs are of petrographic thin sections in plane polarized light. *a*, A photomontage. Scale bars, 10 μm.

trichomes, while the narrower ones probably represent filamentous bacteria or oscillatoriacean algae.

Although septate filaments are rare and occur in a poor state of preservation, one nearly complete specimen (Fig. 3*a*) has been found. The filament is solitary and has a straight to slightly curved shape. The trichome composed of disk-shaped cells 9.0–11.0 μm in diameter and 3.5–5.5 μm in length is uniseriate, unbranched, not constricted or slightly constricted at septa and gradually tapering towards apices. The sheath (Fig. 3*a*, arrow) is hyaline, nonlamellated and up to 7.5 μm thick, gradually tapering towards apical portion. This specimen seems comparable with oscillatoriacean blue-green algae. Oscillatoriaceae-like filaments are very common and important among Precambrian microfossils, which have been found throughout the world[2-7]. The present form constitutes the first reliable record of them in the Precambrian rocks of China.

Sedimentological studies indicate that the Doushantuo Formation represents shallow water marine sediments formed in an arid climate[1]. The occurrence of the new Doushantuo assemblage predominantly composed of coccoid and filamentous cyanophycean algae supports this view. These microfossils seem comparable with some stromatolite-building organisms found in many Precambrian formations[2,5,8-11]. However, in China although rich microfossil assemblages and stromatolites have been described from Precambrian rocks, our knowledge about stromatolite-building organisms in these rocks remains very meagre, and few stromatolites from the Doushantuo Formation of South China have been reported[12]. On the other hand, this newly discovered microbiota represents one of the youngest Precambrian microfossil assemblages now known; others include those from the Vendian of South Kazakstan[10,13], the Arabian Shield[14] and the Lower Yudoma Suite[15].

I thank Professors J. Sutton and J. Watson for encouragement, Professor D. J. Shearman, Dr C. J. Peat and Dr W. L. Diver for valuable discussions, P. Grant, Huang Zhicheng and Ye Jun for rock identification, and Zhang Junming for collecting the rock sample.

Received 29 October 1980; accepted 9 January 1981.

1. Zhao Ziqiang, Xing Yusheng, Ma Guogan, Yu Wen & Wang Ziqiang in *Research on Precambrian Geology: Sinian Suberathem* (ed. Tianjin Institute of Geology and Mineral Resources, Chinese Academy of Geological Sciences), 31 (Tianjin Science and Technology Press, Tianjin, 1980).
2. Schopf, J. W. *J. Paleontol.* **42,** 651 (1968).
3. Oehler, D. Z. *Alcheringa* **2,** 269 (1978).
4. Horodyski, R. J. & Donaldson, J. A. *Precambr. Res.* **11,** 125 (1980).
5. Awramik, S. M. & Barghoorn, E. S. *Precambr. Res.* **5,** 121 (1977).
6. Peat, C. J., Muir, M. D., Plumb, K. A., McKirdy, D. M. & Norvick, M. S. *BMR J. Aust. Geol. Geophys.* **3,** 1 (1978).
7. Maithy, P. K. & Shukla, M. *Palaeobotanist* **23,** 176 (1977).
8. Hofmann, H. J. *J. Paleontol.* **50,** 1040 (1976).
9. Schopf, J. W. *et al. Precambr. Res.* **4,** 269 (1977).
10. Schopf, J. W. & Sovietov, Yu. K. *Science* **193,** 143 (1976).
11. Diver, W. L. *Nature* **247,** 361 (1974).
12. Cao Ruiji & Zhao Wenjie *Mem. Nanjing Inst. Geol. Palaeont. Acad. Sinica* **10,** 1 (1978).
13. Schopf, J. W. *Precambrian Res.* **5,** 143 (1977).
14. Cloud, P., Awramik, S. M., Morrison, K. & Hadley, D. G. *Precambrian Res.* **10,** 73 (1979).
15. Lo, S. C. *Precambr. Res.* **13,** 109 (1980).

Asymmetrical competition in insects

J. H. Lawton

Department of Biology, University of York, Heslington,
York YO1 5DD, UK

M. P. Hassell

Imperial College at Silwood Park, Ascot, Berks SL5 7PY, UK

It is widely believed that when populations of two species of animals compete, each has an adverse effect on the other[1,2]. In other words, each zero growth isocline ($dN_i/dt = 0$ where N_i is the population density of species i) is assumed to be some negative function of the population density of the other species. Cases where one species has a marked effect on the other, but there is no detectable reciprocal effect are sometimes distinguished as 'amensalism'[2,3]. This is regarded as unusual. Standard textbooks of ecology either do not mention it[1,4-6], define it but do not discuss it[2], or occasionally give a brief account[3]. Ricklefs[7] states that "competition can manifest itself by reducing the numbers of one or both competing species" without saying which is the more usual. Here we show that for insects in natural conditions, strongly asymmetrical competition (amensalism or near amensalism) is the norm rather than the exception by a ratio of at least 2:1.

0028-0836/81/090793—03$01.00

Table 1 Examples of amensalism (strongly asymmetrical competition) between insect populations in the field

Ref.	Dominant (j)	Subordinate (i)	Mechanism	Taxonomic group, habitat and food
26	*Calliphora auger*	*Lucilia cuprina*	Resource exploitation and possible habitat modification	Diptera: sheep carrion
26	*Chrysomyia rufifacies*	*Lucilia cuprina*	Resource exploitation and aggression	Diptera: sheep carrion
28	'Early' Odonata	'Late' Odonata	Predation	Odonata: aquatic predators
30*	*Danaus plexippus*	*Oncopeltus* spp.	Extreme exploitation of food-plant	Lepidoptera and Hemiptera: milkweed plants
23	*Bombus pratorum*	*Bombus agrorum*	Aggression	Hymenoptera: artificial, experimental flowers
27†	*Sigara macropala*	*Hesperocorixa lobata*	Exploitation?	Hemiptera: aquatic detritivores
40	*Dacus dorsalis*	*Ceratitis capitata*	Larval interactions; mechanism?	Diptera: guave fruits
22	*Prenolepis imparis*	*Aphaenogaster rudis*	Aggression	Hymenoptera: experimental protein baits in woodland
41*	Several ant species	*Macromischoides aculeatus*	Aggression?	Hymenoptera: predators in cocoa plantations
42	*Fiorinia externa*	*Tsugaspidiotus tsugae*	Resource exploitation and action of shared parasitoid	Homoptera: eastern hemlock foliage
29	*Cromaphis juglandicola*	*Panaphis juglandis*	Habitat modification via excretion	Homoptera: walnut foliage
24	*Bombus terricola*	*Bombus ternarius*	Behavioural avoidance by subordinate	Hymenoptera: goldenrod flowers
25	'Mordellidae 23'	*Epiblema* sp.	Aggression and death of subordinate	Coleoptera and Lepidoptera: plant stems in prairie
43	*Copestylum* cf. *obscurior*	*Quichuana picadoi*	?	Diptera: water-filled *Heliconia* bracts: nectar, detritus and plant tissue
43	*Gillisius* sp.	*Beebeomyia* sp.	?	Coleoptera and Diptera: water-filled *Heliconia* bracts: plant tissue
43	*Gillisius* sp.	*Cephaloleia puncticollis*	?	Coleoptera: water-filled *Heliconia* bracts: plant tissue
43	*Merosargus* sp.	*Gillisius* sp.	?	Diptera and Coleoptera: water-filled *Heliconia* bracts: detritus and plant tissue
43†	*Quichuana picadoi*	*Gillisius* sp.	?	Diptera and Coleoptera: water-filled *Heliconia* bracts: detritus and plant tissue
44	*Copestylum roraima*	*Cephaloleia neglecta*	Disturbance and behavioural avoidance	Diptera and Coleoptera: water-filled *Heliconia* bracts: detritus and plant tissue
45	*Cephaloleia neglecta*	*Xenarescus monocerus*	Interference leading to emigration	Coleoptera: *Heliconia* bracts
18	*Urophora solstitialis*	*Rhinocyllus conicus*	Habitat modification via gall formation	Diptera and Coleoptera: plant tissue in flowerheads of thistle
18	*Larinus sturnus*	*Rhinocyllus conicus*	Probably exploitation of resources	Coleoptera: plant tissue in flowerheads of thistle
18	*Encosma* and *Homoeosomo*	All other species in flower heads	Aggression and predation	Lepidoptera and other: plant tissue in flowerheads of thistle

The dominant species, j, has a marked effect on the subordinate, i, ($\alpha_{ij} \gg 0$) whilst the reciprocal effect is negligible ($\alpha_{ji} \simeq 0$). Less certain, probably amensal, examples are marked *. There are two cases (†) in which the inferior species (i) seems to have a beneficial effect on j, moving the zero growth isocline of the latter clockwise, to intersect the N_1 axis of Fig. 1*b* at >90°.

The view that interspecific competition plays a key part in structuring guilds of birds and mammals is central to current thinking in community ecology[8,9]. Insects have been relatively neglected in this context, despite being the most important terrestrial, and amongst the most important freshwater, consumers. Insects have vastly more species[10] and frequently dominate terrestrial communities in terms of biomass and energy transfer. Widespread amensalism amongst insects would thus have important implications for theoretical community ecology.

Having reviewed the evidence for interspecific competition amonst insect populations in the field[11], we have found major and consistent differences between groups in the apparent importance of competition. Thus, whilst the ants and bees provide many good examples of competition, competition for food between phytophagous species is unusual[12]. Where species do compete, irrespective of their taxonomic group or the type of resource exploited, competition is generally highly asymmetrical. The 23 such examples that we have found are listed in Table 1; they probably provide a conservative total as some involve more than one pair of species. Only those cases where we can be reasonably certain of the degree of asymmetry in the interaction have been included. There are several other reports in which the effects of one species on another are clearly established, but the magnitude of the reciprocal interactions is not given, or cannot be deduced with any degree of certainty (see refs 13, 14 for example). Note that the criterion for amensalism is not simply that one species excludes another: this may happen when both α_{ij} and $\alpha_{ji} > 0$. Instead we require that $\alpha_{ij} > 0$ but $\alpha_{ji} \simeq 0$ (Fig. 1).

We have found only five reports of unequivocal, reciprocal competition involving insect populations in the field—four species-pairs and one larger group. Two species of leaf-feeding Homoptera in the genus *Eupteryx* compete reciprocally[15], as do the bees *Bombus appositus* and *Bombus flavifrons*[16], the aquatic Hemiptera *Arctocorisa carinata* and *Callicorixa producta*[17], and a pair of parasitic Hymenoptera in the genus *Eurytoma*[18]. In addition, McClure and Price[19] demonstrate competition in eight pair-wise experiments involving seven species of Homoptera all in the genus *Erythroneura.* If all eight pairs of interactions show statistically significant reciprocal effects (it is not clear from their data whether this is so), the ratio of examples showing amensalism (Table 1) to those showing more symmetrical competition is approximately 2:1 (23:12). If we choose to count the intensely studied genus *Erythroneura* only once, this ratio rises to nearly 5:1 (23:5). Whatever the exact ratio, the notion that amensalism is unusual is clearly not supported by the data: quite the reverse is true.

Obviously, the data must be interpreted carefully. We have probably overlooked some examples and may have misinterpreted others. Omission of uncertain but possibly symmetrical examples is balanced by cases where amensalism is probable but not proved[20]. Similarly, although there are cases of partial behavioural dominance, and hence possible reciprocal competition (for example, the ants *Pogonomyrmex rugosus* and *Pogonomyrmex desertorum*[21]), these are balanced by examples of almost certain total dominance (*Camponotus* ants frequently kill, maim and displace *Aphaenogaster* in confrontation at food sources[22]). Complete exclusion from preferred resources[22–24] undoubtedly affects colony dynamics in a highly asymmetrical manner. In brief, we think it unlikely that enough examples have been overlooked or misinterpreted to alter our conclusions significantly.

Given the difficulties in determining levels of competition in field conditions and the inevitable noise in the data, it is, of course, impossible to distinguish cases of true amensalism ($\alpha_{ij} > 0$; $\alpha_{ji} = 0$) from cases which are nearly so ($\alpha_{ji} \simeq 0$), and there is undoubtedly a continuum of examples from the truly symmetrical ($\alpha_{ij} \simeq \alpha_{ji} > 0$) to complete amensalism. However, the evidence suggests that examples are not uniformly distributed along this continuum. Strongly asymmetrical competition seems to be the general case, rather than the exception.

The examples of strongly asymmetrical competition summarized in Table 1 cover a wide range of feeding behaviour, including phytophages[25], carrion feeders[26], aquatic detritivores[27], predators[28] and nectivores[24]. They also arise in several very different ways, involving simple avoidance behaviour[24] of one species by another, direct aggression[22] (including, in extreme cases, the death of the inferior competitor[25]), habitat modification[29] or simple exploitation of resources[30]. In several cases the mechanism is not known with certainty; only its effect. The breadth of these examples, and the range of habitats and taxa involved, strengthen our belief that strongly asymmetrical competition is the norm in insects.

Strongly asymmetrical competition approaching amensalism has been widely reported in other groups of animals and, just as for insects, may well be much commoner in these than has generally been acknowledged. Examples include organisms as different as barnacles[31], hummingbirds[32], small mammals[33] and wildebeest and buffalo[34]. Morse[35] provides an extensive review of social dominance between animal species (mainly mammals, birds and fish, but also nine insect examples and one involving Crustacea). Interactions are usually high asymmetrical, particularly where the fundamental niche of an aggressively dominant species is 'included' within that of a subordinate species[35,36]. We emphasize, however, that our examples include several forms of competition, not just social dominance.

We expect the dynamics of species associations in which amensalism is the norm to be very different from those dominated by symmetrical interactions. For example, predictions about niche shifts[32], stability[37,38] and estimates of connectance[39] are all likely to be affected. We believe that studies of community organization for those groups of species in which competition is known to be important should now pay much more attention to highly asymmetrical interactions, treating amensalism as a common, widespread form of species interaction.

We thank J. D. Currey, M. C. Crawley, J. Soberon, J. Waage and M. H. Williamson for valuable comments on the manuscript.

Received 2 October, accepted 16 December 1980

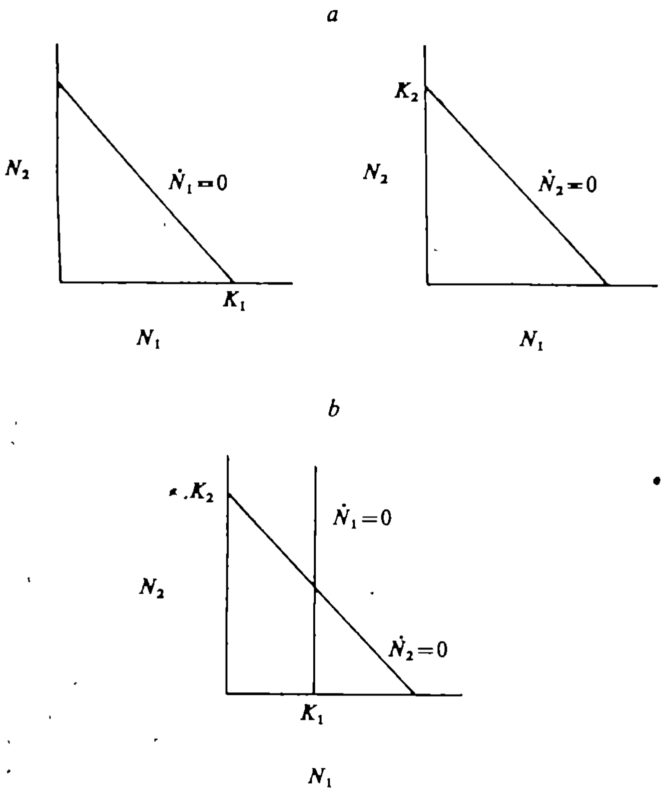


Fig. 1 *a*, Zero growth isoclines for the conventional two-species Lotka–Volterra competition model[1,3].

$$\frac{dN_1}{dt} = r_1N_1(1 - N_1/K_1 - \alpha_{12}N_2/K_1)$$

$$\frac{dN_2}{dt} = r_2N_2(1 - N_2/K_2 - \alpha_{21}N_1/K_2)$$

N refers to the population densities of species 1 and 2, *r* to their intrinsic rates of increase, *K* to their carrying capacities, and α_{12} and α_{21}, the competition coefficients, to the competitive effects of one species on the other. Stable coexistence or the exclusion of one species by another depends on the relative position of the two isoclines. *b*, With amensalism, the population of one species (species 2 in this example) is markedly affected by the other (that is, $\alpha_{21} > 0$), but the reciprocal effect is zero ($\alpha_{12} = 0$). If the isoclines intersect, as shown here, both species coexist. Otherwise the affected species (species 2) is inevitably excluded.

1. Krebs, C. J. *Ecology. The Experimental Analysis of Distribution and Abundance* (Harper & Row, New York, 1978).
2 Odum, E. P. *Fundamentals of Ecology* (Saunders, Philadelphia, 1971).
3. Williamson, M. *The Analysis of Biological Populations* (Edward Arnold, London, 1972).
4. Colinvaux, P. A. *Introduction to Ecology* (Wiley, New York, 1973)
5 Collier, B D., Cox, G W., Johnson, A. W. & Miller, P C *Dynamic Ecology* (Prentice-Hall, London, 1974).
6 Whittaker, R. H *Communities and Ecosystems* (Macmillan, New York, 1975)
7 Ricklefs, R E *Ecology* (Thomas Nelson, London, 1973)
8. Cody, M L & Diamond, J M. (eds) *Ecology and Evolution of Communities* (Belknap, Cambridge, Massachusetts, 1975).
9 Hutchinson, G A *An Introduction to Population Ecology* (Yale University Press, Cambridge, Massachusetts, 1975)
10. Mound, L A & Waloff, N (eds) *Diversity of Insect Faunas* (Blackwell, Oxford, 1978).
11. Lawton, J H & Hassell, M P. in *Ecological Entomology* (ed. Huffaker, C. B.) (in the press)
12 Lawton, J. H. & Strong, D R Jr *Am Nat* **117** (in the press).
13 Pontin, A. J. *J Anim Ecol.* **38**, 747–754 (1969)
14 Force, D C *Science* **184**, 624–632 (1974).
15 Stiling, P. D. *J Anim. Ecol* **49**, 793–805 (1980).
16. Inouye, D W *Ecology* **59**, 672–678 (1978)
17. Pajunen, V I. *Ann Zool Fenn.* **16**, 195–200 (1979)
18. Zwölfer, H *Fortschr. Zool* **25**, 331–353 (1979)
19 McClure, M S. & Price, P W *Ecology* **56**, 1385–1397 (1975).
20. Istock C A. *Ecology* **48**, 929–937 (1967)
21. Davidson, D. W. *Am Nat.* **116**, 92–105 (1980).
22 Lynch, J F., Vail, S. G. & Balinsky, E C *Ecol Ent* **5**, 353–371 (1980)
23. Brian, A. D. *J. Anim. Ecol* **26**, 71–98 (1957)
24 Morse, D H *Science* **197**, 678–680 (1977).
25. Rathcke, B J. *Ecology* **57**, 76–87 (1976).
26. Andrewartha, H. G. & Birch, L. C. *The Distribution and Abundance of Animals* (University of Chicago Press, 1964)
27 Istock, C. A. *Am Nat* **111**, 279–287 (1977).
28 Benke, A. C. *J Anim Ecol* **47**, 335–350 (1978).
29 Messenger, P S in *Insects, Science and Society* (ed Pimentel, D) (Academic, New York, 1975)
30 Blakley, N R & Dingle, H *Oecologia* **37**, 133–136 (1978)
31 Connell, J. H *Ecology* **42**, 710–723 (1971).
32 Pimm, S L *Am Zool* **18**, 797–808 (1978)
33 Crowell, K L. & Pimm, S. L *Oikos* **27**, 251–258 (1976)
34 Sinclair, A. R E *The African Buffalo* (University of Chicago Press, 1977)
35 Morse, D H *Am. Nat.* **108**, 818–830 (1974).
36 Miller, R. S *Adv ecol Res* **4**, 1–74 (1967).
37 Strobeck, C *Ecology* **54**, 650–654 (1973)
38 Gilpin, M. E. & Case, T. J. *Nature* **261**, 40–42 (1977)
39. Rejmánek, M & Starý, P. *Nature* **280**, 311–313 (1979)
40 Kemer, L *et al* *J econ Ent* **67**, 355–360 (1974).
41 Mayer J D *J appl Ecol* **13**, 123–144 (1976)
42 McClure, M S *Ecol Ent* **6**, 47–54 (1981)
43. Seifert, R P & Seifert, F H *Am. Nat* **110**, 461–483 (1976)
44 Seifert, R. P & Seifert, F. H. *Ecology* **60**, 462–467 (1979)
45 Seifert, R P. & Seifert, F H. *Biotropica* **11**, 51–59 (1979).

Time course of O^6-methylguanine removal from DNA of *N*-methyl-*N*-nitrosourea-treated human fibroblasts

A. S. C. Medcalf* & P. D. Lawley

Pollards Wood Research Station, Institute of Cancer Research,
Nightingales Lane, Chalfont St Giles,
Buckinghamshire HP8 4SP, UK

Methylating mutagens are thought to act by induction of direct miscoding, mainly through formation of O^6-methylguanine (O^6MeG) in DNA of, for example, T-even bacteriophage[1], *Escherichia coli*[2], and a cultured Chinese hamster cell line[3]. A positive correlation between this reaction in DNA of target tissues and subsequent induction of tumours has been found for rat kidney tumours[4] and thymic lymphomas in mice[5]. These effects are thought to be counteracted by an enzymatic repair process, which removes O^6-MeG from DNA (first found in *E. coli*[6], then in liver of rats[7] and mice[5]). The predominance of liver as the site of this action *in vivo* led to suggestions[7] that this could account for the relative lack of susceptibility of liver tissue to methylating carcinogens, and also that the enzyme involved might be inducible[8]. Evidence for the latter has been claimed for rat liver[9], but is most convincing in *E. coli*[10] where inhibition of this enzymatic action is achieved by treating cells with chloramphenicol before methylation[11], thus specifically inhibiting protein synthesis. As there is very limited removal of O^6-MeG from methylated cellular DNA, compared with that of 3-methyladenine (3-MeA, a suggested DNA template-inactivating base), it is thought that the former reaction is mediated by a protein which is consumed by the reaction[12]. There have been reports[13,14] that human cells also can remove O^6-MeG, but the data did not clarify the time- and dose-dependencies of the process. We report here more effective removal at lower extents of methylation, which suggests a nonlinear dose response to methylating mutagens and carcinogens for human cells. There were no significant deficiencies in methylpurine removal from DNA of patients with ataxia telangectasia (AT) or xeroderma pigmentosum (XP).

We used untransformed rather than transformed cells as the latter might be relatively deficient in DNA repair capability; density-inhibited cultures[15] were used as it was considered necessary to treat the different cell strains in identical conditions. This eliminated possible cell cycle effects and also DNA synthesis as complicating factors. The cells used were 1BR2, 2BI (radiation repair-proficient), AT3BI and XP12BE (radiation repair-deficient.)

Human fibroblasts were treated with the methylating carcinogen *N*-[^{14}C]methyl-*N*-nitrosourea (MNUA) at doses of ~0.1 mM and ~0.5 mM and the amounts of ^{14}C-methylated purines in cellular DNA were determined after various times of incubation. 3-MeA was rapidly removed from DNA at both doses with a half life of ~2 h; the half life of 7-MeG was ~30 h. These rates are much greater than for the spontaneous hydrolytic loss of these purines from DNA and therefore indicate enzymatic removal processes. The chemically stable O^6-MeG was also removed from DNA (possibly by demethylation as in *E. coli*[16]) but the kinetics differed from those for the other methylpurines—a rapid component was found early on which removed ~2–4 μmol O^6-MeG per mol DNA-P (~4–8 × 10^4 molecules per cell) in the first 1–2 h, then the rate decreased to about one-fifth of its initial value. At the lower doses of MNUA, 90% of the O^6-MeG was removed within 10 h, whereas at the higher dose level over half still remained. Regarding the hypothesis that O^6-MeG is a directly miscoding base, which causes mutation and tumour initiation if not removed before replication of methylated DNA templates, more effective removal at lower extents of methylation suggests a nonlinear dose response.

The results are summarized in Fig. 1 and Table 1, presented respectively as amounts of methylated bases per 10^6 DNA nucleotides, and as ratios of methylated bases at various incubation times after methylation. No corrections were required for dilution of bases by DNA synthesis, as there was no significant reduction of specific radioactivity of ^{3}H-thymidine pre-label in cellular DNA. Clearly the methylpurines examined were actively removed from DNA, 3-MeA and 7-MeG much more rapidly then occurs by spontaneous hydrolysis (half lives of ~30 and ~110 h respectively)[17].

Figure 1 shows that the time course of removal of each methylpurine from DNA had its own specific characteristics. 7-MeG declined with approximately first-order kinetics (linear semi-logarithmic plot) and with the same half life at both dose levels. 3-MeA was removed much more rapidly, with nonlinear kinetics, but again the efficiency of removal was independent of dose. However, O^6-MeG showed markedly nonlinear kinetics, and approximated to a two-component time course, with a higher efficiency of removal at the lower dose than at the higher.

This conclusion rests on the assumption that the initial extents of methylation of cellular DNA at the O-6 atom of guanine can be calculated as 0.11 times the extent of methylation at the N-7 atom, that is, they are identical to those found for methylation of DNA in the isolated state *in vitro*, or in cells which cannot remove this base[3]. These calculated values are indicated by arrows on the *y* axes of Fig. 1. The extrapolations of the approximately linear portions of the graphs for the higher doses (Figs 1*b*, *d*) to zero time are quite close to, but significantly lower than, these calculated values. The earliest time points for which analyses of DNA could be carried out were 1–2 h after addition of the methylating agent, to allow for completion of the methylation reaction (half life ~0.2 h) and collection of cells for DNA isolation.

During this time, as is evident from the time courses for removal of O^6-MeG (Fig. 1) ~2 μmol per mol DNA-P atoms (or DNA nucleotide units) were removed at the lower dose, and ~3–4 μmol per mol DNA-P at the higher dose—this represents a much higher proportion of the initial reaction at the lower (~3 μmol per mol DNA-P) than at the higher dose (15–20 μmol per mol DNA-P). It was calculated from the known DNA content of the cells (~2 × 10^{10} DNA-P) that this rapid initial component removed 4–8 × 10^4 molecules O^6-MeG per cell per h, or ~1,000 per cell per min.

The rate of removal then declined at both dose levels, the linear portion of the graphs at the higher dose showing a rate of ~200 molecules per cell per min. These kinetic characteristics could be considered to support the view[12] that the initial process is 'quasi-enzymic', that is, that the protein involved is consumed by the reaction, and is initially present in a limited amount.

In view of reports that in *E. coli* this enzyme (or quasi-enzyme) is inducible by pretreatment of cells with low doses of methylating agents[12], the question arises of whether the enzyme in human cells is similarly inducible. Evidence for this is not yet available, but the observed rapidity of the initial component reaction in the present work contraindicates its induction by the experimental treatment. This follows from the finding that removal of O^6-MeG can be at least as rapid in human cells as that of 3-MeA, whereas in *E. coli* removal of the latter, which is effected by a constitutive enzyme, is more rapid, and enzyme induction in mammalian cells is expected to be a much slower process than in bacteria[18]. It is possible that, due to low levels of methylating carcinogens in the diet (for example, dimethylnitrosamine in meat products[19]), the enzymatic activity in isolated cells could have been preinduced in the donors. It might be difficult to distinguish between such preinduced activity and constitutive activity, but the persistence of preinduced activity through many cell generations in culture seems improbable.

---

* Present address: Carcinogenesis Laboratory, Fee Hall, Michigan State University, East Lansing, Michigan 48824, USA.

0028–0836/81/090796—03$01.00

Table 1 Ratios of methylated bases at various incubation times after methylation

Cell line	MNUA dose (mM)	Time (h)	% Survival	O^6-MeG/7-MeG	3-MeA/7-MeG
1BR2	0.11	1.2	90.0	0.036	0.061
		5.5	—	0.011	0.005
	0.46	1.2	14.8	0.077	0.053
		4.0	—	0.070	0.015
		23.0	50.5	0.079	—
AT3BI	0.09	1.2	—	0.050	0.044
		4.0	—	0.019	0.035
		20.0	—	0.005	0.005
	0.50	1.0	1	0.090	0.049
		4.0	—	0.091	0.022
		21.0	38.3	0.090	0.008
2BI	0.18	1.8	—	0.035	0.061
	0.51	1.1	13.0	0.082	0.058
		2.8	—	0.089	0.031
		24.0	58.7	0.051	—
XP12BE	0.10	1.2	—	0.032	0.050
		28.0	—	0.006	0.018

Values of extents of methylation of DNA were obtained as described in the legend to Fig. 1, and additional experiments have confirmed the results with 1BR2, AT3BI and XP12BE. Survivals were determined by collecting density-inhibited cultures by trypsinization, and plating appropriate numbers of cells to allow for colony growth at plating efficiencies of 10–40%, using feeder layers as described elsewhere[24]. Survival is colony-forming ability of treated cells divided by that of controls, only colonies with >50 cells being scored. Additional experiments (data not shown) were done to show that survival in all cases increased to almost 100% after ~40 h incubation in the density-inhibited state.

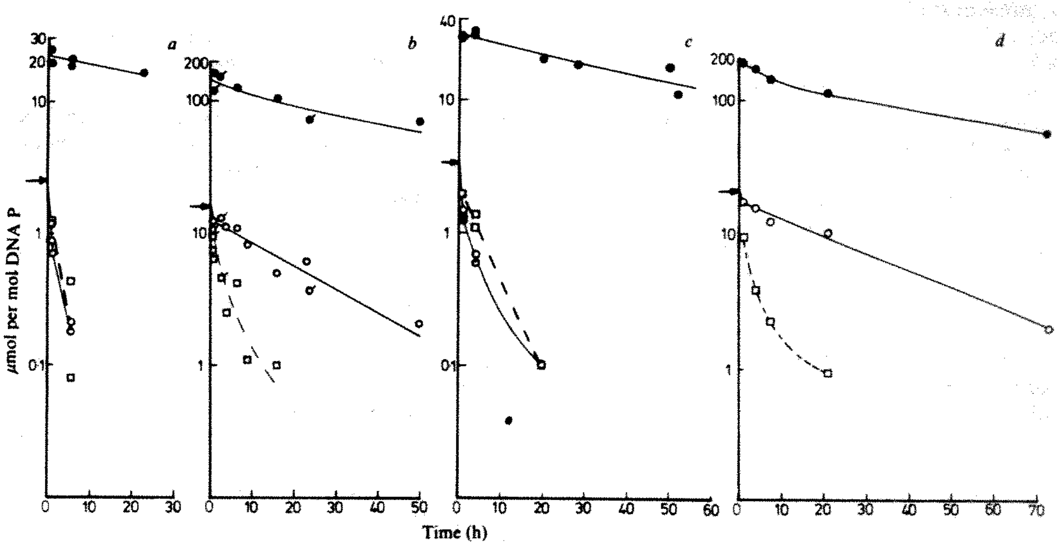


Fig. 1 These experiments were carried out in as near similar conditions as possible to eliminate the effects of variables such as DNA synthesis. The use of confluent cells was considered the most appropriate and convenient way of achieving this. Cells were obtained in this condition by a modification of the method of Weichselbaum *et al.*[15]. Cells (2×10^6) were inoculated into plastic tissue culture flasks (Nunc, 175 cm^2) containing Dulbecco's modified minimal essential medium (50 ml) supplemented with 15% fetal calf serum (Gibco). After incubation at 37 °C (48 h), ^3H-thymidine (Radiochemical Centre) was added (3 µCi $mmol^{-1}$, 7.4 µM), and the cells were collected after 24 h using trypsin/EDTA, pooled and the required number of flasks seeded with 2×10^6 cells in 50 ml medium; this ensures that there is equal specific radioactivity of ^3H-thymidine in each flask. There were only negligible decreases in the specific radioactivities of ^3H-thymidine in DNA over the experimental period, as expected for density-inhibited cells. The flasks were further incubated at 37 °C until confluence was reached, when the medium was collected for use as conditioned medium. This was repeated twice at daily intervals and the experiment was carried out 24 h after the last re-feed. Medium was removed from each flask, and the cells were washed once in prewarmed phosphate-buffered saline (PBS), residual PBS being removed by pipette. ^{14}C-MNUA (Radiochemical Centre, 37 mCi $mmol^{-1}$) was added to prewarmed PBS to give the required concentration for treatment, from a 1:200 volume of a stock solution in anhydrous dimethyl sulphoxide (DMSO), and 10 ml was added to each flask. The concentration was checked by assay of the radioactivity; controls were treated with DMSO (0.5% v/v). This concentration of DMSO was non-toxic and experiments with 1% (v/v) or no DMSO gave identical results. Treatment was for 40 min at 37 °C in the dark and was terminated by removal of radioactive PBS, and washing the cells once with warm PBS. At this stage and at appropriate times after incubation in conditioned medium, cells were collected from 1–4 flasks (dependent on the dose used). Sufficient cells were retained for assay of survival by the method of Cox and Masson[24]. The remaining cell residue ($0.3-1.5 \times 10^8$ cells) was used for isolation of DNA by a modification of the method of Kirby[25]. DNA (0.2–1 mg) was hydrolysed in 0.1 M HCl at 70 °C to liberate purine bases and analysed chromatographically using HPLC on a Partisil 10-SCX column eluted with a gradient of 0.05–0.4 M ammonium formate, 12–10% methanol, essentially as described by Frei *et al.*[5] Guanine and adenine were determined in each chromatogram by UV absorption of the appropriate fractions: ε_{max}(guanine) 8,000 at 273 nm; ε_{max}(adenine) 13,100 at 260 nm. The amount of DNA-P was then calculated as 2 × (mol guanine + mol adenine). This gave the extents of methylation of DNA, principally at O-6 and N-7 of guanine and N-3 of adenine. The initial extents of reaction at O-6 of guanine were calculated as 0.11 times the values for 7-MeG and are shown by arrows. *a*, 1BR2 treated with MNUA, (~0.1 mM); *b*, 1BR2 and 2BI (denoted by ○, □ and ●) treated with MNUA (~0.5 mM); *c*, AT3BI treated with MNUA (~0.1 mM); *d*, AT3BI treated with MNUA (~0.5 mM). ○, O^6-MeG; □, 3-MeA; ●, 7-MeG.

However, the slower component of the removal could be induced by the methylation damage to DNA.

Regarding a possible association between inherited defects in repair capability for radiation-induced damage in DNA and defective ability to remove methylpurines, our results showed no evidence for such an association in the AT or XP cells used. Previous reports[13,14] which claimed XP-associated defective repair of methylation damage were based on the use of virally transformed XP cells; therefore it cannot be ruled out that the defective removal of O^6-MeG was associated with the transformed nature of the cells, and not with any defect in UV repair capability.

Several implications for consideration of dose–response relationships in the associated mutagenic, transforming and tumour initiating effects of methylating carcinogens can be made from our results. Such relationships could explain why human cells have proved refractory to malignant transformation *in vitro* by such agents. One study[20] concluded that this transformation, by *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine required a relatively high degree of toxic action of the methylating agent, killing about 70% of the cells, that is, a correspondingly high extent of methylation of cellular DNA, possibly sufficient to reduce markedly the efficiency of O^6-MeG removal.

The levels of methylating carcinogens in the human diet are probably low[19] (for example, ~1–10 parts per 10^9 of dimethylnitrosamine (DMNA) was estimated in certain meat products) and it seems likely that these would lead to sufficiently low extents of DNA methylation to be removed very efficiently in cells with the same ability to carry out this reaction as the fibroblasts of the present study.

Analogous ability of cells of rat liver to remove this miscoding base from DNA, methylated by continuous feeding with DMNA, has been suggested[7] to account for the relatively low (or, on the concept that a no-effect threshold dose exists, zero) yields of induced liver cell cancer at lower dose levels (suggested threshold of 0.1 parts per 10^6)[21].

It is also relevant to consider the frequency of UV-induced mutation in human fibroblasts, which shows a no-effect threshold[22]. This threshold is associated both with efficient repair of UV-induced DNA damage, and with protection against the induction of mutation and cancer, because XP cells, deficient in this repair, show no threshold, and are more prone to the mutagenic and carcinogenic effect of UV light.

Another case of a markedly nonlinear dose response has been reported[23] for induction of mutation in human lymphoblasts by a methylating carcinogen, *N*-methyl-*N*-nitrosourethane, and this could also be ascribed to more efficient removal of the promutagenic base O^6-MeG at low doses.

Taken together, these considerations lead to the general hypothesis that cells with the ability to remove O^6-MeG from DNA more efficiently at lower extents of methylation, will show a markedly nonlinear dose response for malignant transformation by methylating carcinogens, with a very low probability per unit dose at low doses. In the present study the human cells examined were in this category and therefore should be correspondingly resistant to low doses of methylating carcinogens.

We thank Dr C. F. Arlett for supplying 1BR2 and 2BI cells, Professor D. G. Harnden for AT3BI cells, and Dr J. J. Roberts for XP12BE cells. This work was supported by grants from the CRC and MRC; A.S.C.M. was a postgraduate research student of the Institute of Cancer Research.

Received 5 June; accepted 25 November 1980.

1. Loveless, A. *Nature* **223,** 206–207 (1969).
2. Jeggo, P. *J. Bact.* **139,** 783–791 (1979).
3. Newbold, R. F., Warren, W., Medcalf, A. S. C. & Amos, J. *Nature* **283,** 596–599 (1980).
4. Nicoll, J. W., Swann, P. F. & Pegg, A. E. *Nature* **254,** 261–262 (1975).
5. Frei, J. V., Swenson, D. H., Warren, W. & Lawley, P. D. *Biochem. J.* **174,** 1031–1044 (1978).
6. Lawley, P. D. & Orr, D. J. *Chemico-biol. Interactions* **2,** 154–157 (1970).
7. O'Connor, P. J., Capps, M. J. & Craig, A. W. *Br. J. Cancer* **27,** 153–166 (1973).
8. Lawley, P. D. *IARC Sci. Publ.* **12,** 181–210 (1976).
9. Montesano, R., Brésil, H. & Margison, G. P. *Cancer Res.* **39,** 1798–1802 (1979).
10. Samson, L. & Cairns, J. *Nature* **267,** 281–282 (1976).
11. Warren, W. & Lawley, P. D. *Carcinogenesis* **1,** 67–78 (1980).
12. Robins, P. & Cairns, J. *Nature* **280,** 74–76 (1976).
13. Goth-Goldstein, R. *Nature* **267,** 81–82 (1977).
14. Altamirano-Dimas, M., Sklar, R. & Strauss, B. *Mutat. Res.* **60,** 197–206 (1979).
15. Weichselbaum, R. R., Nove, J. & Little, J. B. *Nature* **271,** 261–262 (1978).
16. Karran, P., Lindahl, T. & Griffin, B. *Nature* **280,** 76–77 (1979).
17. Lawley, P. D. & Warren, W. *Chemico-biol. Interactions* **12,** 211–220 (1976).
18. Lehninger, A. L. *Biochemistry* 2nd edn, 832–833 (Worth, New York, 1976).
19. Preussmann, R. in *Primary Liver Tumours* (eds Remmer, H., Bolt, H. M., Bannasch, P. & Popper, H.) 11–29 (MTP Press Ltd., Lancaster, 1978).
20. Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1334–1338 (1978).
21. Arai, M. *et al. Gann* **70,** 549–558 (1979).
22. Maher, V. M., Dorney, D. J., Mendrala, A. L., Konze-Thomas, B. & McCormick, J. J. *Mutat. Res.* **62,** 311–323 (1979).
23. Penman, B. W., Hoppe, H. IV & Thilly, W. G. *J. natn. Cancer Inst.* **63,** 903–907 (1979).
24. Cox, R. & Masson, W. K. *Int. J. Radiat. Biol.* **26,** 193–196 (1974).
25. Kirby, K. S. *Biochem. J.* **66,** 495–504 (1957).

Distinct lung-colonizing and lung-metastasizing cell populations in B16 mouse melanoma

Christopher W. Stackpole*

Memorial Sloan-Kettering Cancer Center, Rye,
New York 10580, USA

Establishment of distant metastases by solid tumours requires that some tumour cells separate from the primary growth, infiltrate surrounding normal tissues, invade blood vessels or lymphatic channels, and extravasate into microenvironments favourable for proliferation[1]. It is probable that only a small subpopulation of tumour cells is endowed with these special attributes[2]. The biological properties of potentially metastatic tumour cells are being investigated extensively by means of an 'experimental metastasis' assay, in which cultured cells are introduced directly into blood vessels and, several weeks later, organs examined for tumour growths[1–8]. Although this assay allows controlled and quantitative experimentation, the suitability of experimental metastasis (colonization) as a valid representation of true metastasis is uncertain. I report here that organ colonization after injection of cells directly into the blood vasculature is not necessarily predictive of spontaneous metastatic potential from subcutaneous transplants. Indeed, some cell lines derived from the B16 mouse melanoma produce numerous lung colonies when injected intravenously (i.v.) but rarely metastasize, whereas others are spontaneously metastatic to the lungs but produce few lung colonies after i.v. injection. These differences in behaviour are evident only when the possibility of colonization occurring due to accidental introduction of melanoma cells into blood vessels is excluded.

B16 melanoma cell lines were maintained as monolayer cultures in 75 cm^2 plastic tissue culture flasks with Eagle's minimal essential medium containing 10% fetal bovine serum, 1.5% vitamins, 1% nonessential amino acids, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units ml^{-1} penicillin, and 100 $\mu g\,ml^{-1}$ streptomycin, *p*H 7.0. Primary lines were established from single-cell suspensions prepared from subcutaneous (s.c.) or intraperitoneal (i.p.) tumour transplants, lung colonies, or lung metastases, by mincing the tissue with scissors and pressing it through 80-mesh stainless-steel wire screens, and washing by low-speed centrifugation, in sterile conditions. Before experiments, monolayers were initiated from frozen stocks by seeding 5×10^4 viable cells and were collected ~24 h before confluence by incubation for 1 min at 37 °C with 1 ml of Hank's balanced salt solution, without calcium or magnesium, supplemented with 0.25% trypsin, 0.02% EDTA and penicillin–streptomycin. Single-cell suspensions free of cell clumps and >95% viable were washed twice with Earle's balanced salt solution containing 1% heparin and penicillin–streptomycin

---

* Present address: Department of Medicine, New York Medical College, Valhalla, New York 10595, USA.

Table 1 Lung colonization by B16-F10 cells injected at various anatomical sites

Injection site	No. of mice with lung colonies	Median no. of lung colonies
Lateral tail vein (i.v.)	20/20	88 (28–250)
Hind leg, footpad (i.m.)	7/20	6 (1–14)
External ear (i.d.)	3/20	12 (9–16)
Hind leg, thigh (s.c.)	5/20	5 (2–10)
Abdominal flank (s.c.)	0/20	0

Female mice 8–10 weeks old were partially immobilized by ether inhalation for 10 s, and 10^5 B16-F10 cultured cells in 0.05 ml of EBSS injected 5 mm proximal to the point of entry of a 27-gauge needle. Injections were into the lateral tail vein 2 cm from the base of the tail; into the middle of the hind foot 2–3 mm beneath the skin on the lower surface after insertion in the footpad; beneath the epidermis of the medial dorsal surface of the ear 5 mm from the ear base; beneath the skin on the ventral surface of the thigh midway between knee and hip; and beneath the skin midway between the front and hind legs 1 cm to the right of the abdomen centre. After 1 h, the injected areas were removed by amputation at least 5 mm proximal to the limit of the inoculum or by excision (flank), including a 5 mm margin beyond the inoculum boundary. Surgery was performed after anaesthetizing mice with chloral hydrate; large blood vessels were ligated and heat-cauterized, and open wounds were resealed with 9-mm stainless-steel wound clips. All mice were killed 21 days later and pigmented lung colonies counted by examining excised lungs with a dissecting microscope; complete autopsies revealed grossly visible tumour growth only in the lungs. These results were obtained in duplicate experiments carried out using 10 mice per group. Values in parentheses are the range of lung colonies in positive mice.

Table 2 Effect of surgical excision on metastasis of B16-F10 s.c. transplants

Procedure	No. of mice with lung nodules	Median no. of lung nodules
2.5×10^4–2.5×10^5 cells s.c. (autopsied at death 4–5 weeks later)	6/95*	8 (1–18)
2.5×10^5 cells s.c. ('careful' complete excision after 14 days; killed 6 weeks later)	1/20	4
2.5×10^5 cells s.c. ('careful' incomplete excision after 14 days; autopsied at death 4–6 weeks later)	11/18	66 (19–160)
2.5×10^5 cells s.c. ('careless' complete excision after 14 days; autopsied or killed 4–6 weeks later)	12/17	53 (5–165)

B16-F10 cultured cells in 0.2 ml of EBSS were injected s.c. in the right abdominal flank as described in Table 1 legend but mice were not anaesthetized before injection. Except as otherwise noted for the first group, mice were females 8–12-weeks old. 'Careful' complete excision was carried out as described in Table 1 legend and involved removal of all visible tumour plus a border of normal tissue at least 5 mm wide without disturbing the tumour; no local tumour growth occurred in any of these mice after excision. Incomplete excision was done in the same manner but only grossly visible tumour was removed; traces of tumour remained to grow and eventually kill all these mice. 'Careless' excision was done without ligating and cauterizing blood vessels, and tumours were handled without caution during removal; local tumour growth occurred in 53% of these mice. The mean size of excised tumours was 1.04 ± 0.45 cm^3, and all were richly vascularized. Values in parentheses are the range of lung nodules in positive mice.

* Included in this group were mice injected with 2.5×10^4 (0/15) and 2.5×10^5 (2/48) cultured cells, and 10^5 tumour cells purified from an i.p. transplant (1/10); male mice 8–12-weeks old injected with 2.5×10^5 cultured cells (1/10), and female mice 12–18-months old injected with 2.5×10^5 cultured cells (2/12).

(EBSS). Cell suspensions in EBSS were injected into disease-free C57BL/6 mice (Jackson Laboratory or our own colonies), matched in sex, age and weight for each experiment, within 10 min after collection[9]. Tumour cells from single-generation s.c. or i.p. transplants were purified to >90% viability and homogeneity by centrifugation on continuous-density gradients of buffered iso-osmotic metrizamide after physical separation of cells from solid tissue[10]. Tumour cells were prepared for injection in the same manner as cultured cells.

The B16-F10 cell line was highly selected by Fidler[3] to produce large numbers of lung colonies after i.v. injection, and s.c., intradermal (i.d.) and intramuscular (i.m.) transplants of these cells are reported to metastasize spontaneously to the lungs in 20–90% of syngeneic mice, usually after excision of the growing transplant[11–14]. With regard to metastatic activity, several of those transplantation sites are richly vascularized, and the possibility of introducing potential lung-colonizing cells directly into ruptured blood vessels during injection must be considered. When B16-F10 cells were introduced into various s.c., i.d. and i.m. sites and tumour prevented from developing at the injection site by amputation or excision of the affected areas, lung colonies developed in some mice (Table 1); in mice with growing transplants such colonies could be mistaken for spontaneous metastases. Only s.c. injection in the abdominal flank, in an area which was predetermined as relatively free of blood vessels visible at low magnification with a dissecting microscope, failed to result in some lung colonization, and this site was chosen for subsequent experiments. Similarly, surgical excision of B16-F10 transplants, while prolonging host survival and growth of pre-established micrometastases, could also allow tumour cells inadvertently to enter severed blood vessels and eventually colonize the lungs—this possibility was also examined. Tumours were excised with great care after 2 weeks of s.c. growth and lung nodules formed in <10% of mice, comparable with the incidence of nodules in mice bearing unexcised tumours (Table 2); these nodules may represent true spontaneous metastases. In contrast, careless or incomplete tumour excision resulted in extensive lung nodule formation, probably due to colonization rather than spontaneous metastasis.

With these considerations, lung colonizing and spontaneous metastatic capacities of a number of cell lines derived from the B16 melanoma were assessed. Each cell line was tested as cultured cells and as tumour cells, because the behaviour of cells growing in such diverse conditions might be different. The 'parent' B16 tumour, from which B16-LC (lung-colonizing) and B16-LM (lung-metastasizing) lines were derived, produced both lung colonies and metastases when tested as tumour cells, but cultured cells were nonmetastatic and only slightly active in colonization (Table 3). Ten B16-LC lines, of which B16-LC1 and B16-LC4 represent extremes in lung-colonizing capacities, and three lines selected by Fidler[3] and Hart[12] (B16-F1, B16-F10 and B16-BL6) were not significantly metastatic (0–20%) either as cultured or tumour cells. Clones isolated from lung-colonizing lines behaved similarly. In contrast, seven B16-LM lines were metastatic in 35–100% of mice. Lines derived from tumour transplants (B16-LM1, -LM2 and -LM3) were both colonizing and metastatic, although cultured B16-LM3 cells showed little activity in either assay (Table 3). Surprisingly, lines obtained from lung metastases (B16-LM4 and B16-LM6) were potently metastatic but at most only poorly colonizing. Twenty-five highly metastatic clones isolated from both types of metastatic line have also been tested—all were poor colonizers. A representative clone (C2) from the B16-LM4 line failed to colonize the lungs when 5×10^3 cultured cells were injected i.v. into five mice, produced few colonies (median 0, range 0–3, $n=5$) when 5×10^4 cells were injected, but formed numerous colonies (median 47, range 19–62, $n=5$) when 5×10^5 cells were injected. Paradoxically, it seems that cell populations that can spontaneously migrate to and grow in the lungs are less successful than many colonizing populations at proliferating in the lungs after direct i.v. introduction.

Table 3 Lung-colonizing and lung-metastasizing cell lines derived from the B16 melanoma

Cell line	Derivation	Growth form tested	No. of mice with lung colonies; median no. (range)	No. of mice with lung metastases; median no. (range)
B16	'Parent' tumour	C	10/10; 2(1–8)	0/10; 0
		T	10/10; 27(10–63)	4/10; 0(0–8)
Lung-colonizing lines				
B16-LC1	s.c. transplant (10)*	C	19/23; 1(0–7)	0/20; 0
		T	10/10; 30(12–66)	1/10; 0(0–3)
B16-LC4	Multiple lung colonies	C	10/10; 250(145–>300)	0/15; 0
		T	10/10; >300(74–>300)	1/10; 0(0–1)
B16-F1	(refs 1–3)	C	17/20; 11(0–25)	1/20; 0(0–1)
		T	10/10; 47(12–76)	0/10; 0
B16-F10	(refs 1–3)	C	25/25; 63(26–116)	1/25; 0(0–3)
		T	10/10; 180(91–>300)	1/10; 0(0–3)
B16-BL6	(refs 12, 14)	C	10/10; 37(16–68)	2/10; 0(0–4)
		T	9/9; 110(63–165)	2/10; 0(0–5)
Lung-metastasizing lines				
B16-LM1	s.c. transplant (18)	C	14/14; 16(4–61)	8/19; 0(0–4)
		T	20/20; 32(4–120)	16/18; 4(0–14)
B16-LM2	i.p. transplant (6)	C	20/20; 20(3–68)	9/24; 0(0–12)
		T	10/10; 75(35–146)	9/10; 6(0–19)
B16-LM3	s.c. transplant (23)	C	4/15; 0(0–6)	2/15; 0(0–1)
		T	14/20; 11(0–82)	14/20; 1(0–16)
B16-LM4	Single lung metastasis	C	5/30; 0(0–7)	26/29; 5(0–17)
		T	10/10; 3(1–15)	9/10; 4(0–26)
B16-LM6	Multiple lung metastases	C	9/20; 0(0–8)	16/19; 4(0–26)
		T	9/10; 4(0–13)	9/9; 4(1–35)

B16-LC and B16-LM lines were derived from transplants, lung colonies and lung metastases originating from the 'parent' B16 melanoma (tumour tissue that, except for brief periods of serial transplantation in syngeneic mice, had been stored frozen since 1962); these lines were tested within two passages after primary isolates were frozen. Other lines (provided by Dr Isaiah J. Fidler of the Frederick Cancer Research Center and by the Mason Research Institute) were also tested within two passages. Female mice 8–12-weeks old were used in all colonization and metastasis assays, which were carried out simultaneously for each cell line. Colonization was assessed by injecting 5×10^4 cells in 0.2 ml of EBSS into the tail vein, then counting lung colonies after 21 days; metastatic activity was determined by injecting 0.25 ml of EBSS containing 2.5×10^5 cultured cells, or 10^5 tumour cells, s.c. in the abdominal flank, then counting lung nodules after 4 weeks of undisturbed tumour growth. Local tumour growth occurred in all mice receiving s.c. injections. C, cultured cells; T, tumour cells.

* Transplant generations are indicated in parentheses.

These results indicate that, based on predominant behavioural characteristics, two distinct cell populations exist in the B16 melanoma: lung-colonizing and lung-metastasizing cells. There is no evidence from analysis of clones to indicate the existence of metastatic cells with potent colonizing abilities. However, the presence of a third population, consisting of noncolonizing nonmetastatic 'null' cells, is indicated by the relatively low activity of cultured B16, B16-LC1 and B16-LM3 cells in colonization and metastasis assays. The considerable behavioural differences between lines derived from the same tumour suggest that the relative proportion of colonizing, metastatic and 'null' populations in the B16 melanoma is highly variable and dependent on growth conditions. This tumour is thus even more heterogeneous and complex than previously thought[15], and the availability of cell lines and clones with such dramatical biological differences from a single tumour should facilitate efforts to elucidate the basis for metastasis.

This study was supported by the General Motors Surgical Research Laboratory of the Memorial Sloan-Kettering Cancer Center and by NIH grant CA-08748. I thank Patricia Stremmel for assistance.

Received 4 September; accepted 8 December 1980.

1. Fidler, I. J., Gersten, D. M. & Hart, I. *Adv. Cancer Res.* **28**, 149–250 (1978).
2. Poste, G. & Fidler, I. J. *Nature* **283**, 139–146 (1980).
3. Fidler, I. J. *Nature new Biol.* **242**, 148–149 (1973).
4. Withers, H. R. & Milus, L. *Cancer Res.* **33**, 1931–1936 (1973).
5. van Putten, L. M., Kram, L. K. J., van Dierendonck, H. H. C., Smink, T. & Füzy, M. *Int. J. Cancer* **15**, 588–595 (1975).
6. Nicolson, G. L., Brunson, K. W. & Fidler, I. J. *Cancer Res.* **38**, 4105–4111 (1978).
7. Stringfellow, D. A. & Fitzpatrick, F. A. *Nature* **282**, 76–78 (1979).
8. Olsen, G. A., Glasgow, L. A. & Dethlefsen, L. A. *Cancer Res.* **40**, 853–860 (1980).
9. Fidler, I. J. *Meth. Cancer Res.* **15**, 399–439 (1978).
10. Stackpole, C. W., Cremona, P. & Kassel, R. L. *J. immun. Meth.* **30**, 231–243 (1979).
11. Fidler, I. J. *Cancer Res.* **35**, 218–224 (1975).
12. Hart, I. R. *Am. J. Path.* **97**, 587–600 (1979).
13. Wang, B. S., McLoughlin, G. A., Richie, J. P. & Mannick, J. A. *Cancer Res.* **40**, 288–290 (1980).
14. Poste, G., Doll, J., Hart, I. R. & Fidler, I. J. *Cancer Res.* **40**, 1636–1644 (1980).
15. Fidler, I. J. *Cancer Res.* **38**, 2651–2660 (1978).

Hypoblastic cells can form a disk inducing an embryonic axis in chick epiblast

E. Mitrani & H. Eyal-Giladi

Embryology Section, Department of Zoology,
The Hebrew University of Jerusalem, Jerusalem, Israel

The primitive streak of the chick embryo develops from one of the two layers of cells of the stage XIII blastoderm[1], the epiblast. The other layer of cells, the hypoblast, seems to be necessary for the induction of the primitive streak and also determines its orientation[2]—rotation of the hypoblast by 90° is followed by a similar rotation of the embryonic axis[3]. After stage XIII, the hypoblast is replaced by the invaginating endoderm and plays no further part in the development of the embryo[4]. By means of a technique for disaggregation and reconstituting cells of stage XIII hypoblasts, we have been able to show that the two functions, induction and orientation are independent and that, with reconstituted hypoblasts, the orientation of the primitive streak is determined by the epiblast.

It was not clear whether the hypoblast's function is just one of dictating the future polarity of the embryo or whether it also performs a specific inductive function. In normal development the hypoblast starts to appear at stage XI. It is formed by cells from two different sources: cells from the posterior part of the marginal zone[5] moving anteriorly, and polyinvaginating cells coming down from the central area of the one-layer blastoderm[1]. At stage XIII, if the fully formed hypoblast is removed, the remaining blastoderm can regenerate a new hypoblast, from the two above distinct sources, which is to generate axial development in the epiblast[6]. If the peripheral regions of the marginal zone and the area opaca are removed at stage XIII, the remaining part will include only the epiblast proper and the hypoblast. These two components are sufficient for development of normal embryos. If from such an operated blastoderm the hypoblast is also removed, the remaining epiblast proper will regenerate a second layer. However, although this layer will be formed in about the same time as when the marginal zone is present during regeneration, it will only be formed from polyinvaginating epiblastic cells and will not be capable of producing axial development in the epiblast. Only unorganized mesoderm such as blood islands will be formed[6]. It seemed possible that in the latter case, the information concerning direction of growth, which is normally transmitted by a normal hypoblast to the epiblast, is lacking, and therefore the epiblast fails to organize the resulting mesoderm in a polarized manner.

We have conceived a model system which allows us to separate induction from polarity within the hypoblast so that the two processes can be studied independently. For this purpose we use a culture system in which individual hypoblastic cells are made to grow and to aggregate into a disk which closely resembles the original hypoblast. This reaggregated hypoblast is then layered on to the ventral surface of a normal stage XIII epiblast and the system is incubated further on a chick vitelline membrane.

In some cases care was taken to mark, on the epiblast, the posterior side of the original blastoderm with carbon particles. The conditions used to obtain appropriate hypoblast aggregates varied. A compromise had to be reached between the minimum time necessary to obtain a proper aggregate and the time at which the functional capacity of the cells is maximal. The best hypoblast aggregates were obtained after culturing the hypoblastic cells for 11 h (Fig. 1*a*). If the hypoblastic cells were cultured for only 5 h the aggregates formed were more difficult to remove as an intact sheet but were as functional as those obtained after 11 h. The results of the experiments in which hypoblast aggregates were layered on to the ventral surface of stage XIII epiblasts are summarized in Table 1. When reconstituted hypoblasts reaggregated for 11 h and layered on top of stage XIII blastoderms deprived of their original hypoblast were incubated further, they were able to induce development of normal embryos. Subsequent experiments were performed by layering the reconstituted hypoblasts on top of stage XIII blastoderms previously deprived of the hypoblast, the marginal zone and the area opaca (Fig. 1*b*, *c*). This ensured that the embryos obtained were not a result of regeneration of a new hypoblast from the marginal zone[6].

In 10 of the experiments performed, the expected posterior side of the embryo was clearly marked in the epiblast. Of these 10, all developed an axis according to the previously marked direction. The degree of development in all the experiments recorded in Table 1 varied from a definite primitive streak to embryos with several pairs of somites.

Table 1 Development and orientation of embryos from epiblasts incubated with reaggregated hypoblasts

	No. of experiments	No. of blastoderms with embryonic axis	Deviation from original blastoderm axis
Epiblasts with marginal zone and area opaca	5	5 (2)	0° (2)
Epiblasts proper, no marginal zone and no area opaca	17	14 (8)	0° (8)

Numbers in parentheses indicate the number of cases in which the direction of the axis was recorded.

Clearly the position of the cells with respect to each other in the primary hypoblast is not critical in bringing about an embryonic axis in the epiblast. The fact that the embryos developed in accordance with the expected axis of the original blastoderm indicates the existence within the epiblast of sufficient information to determine the future direction of

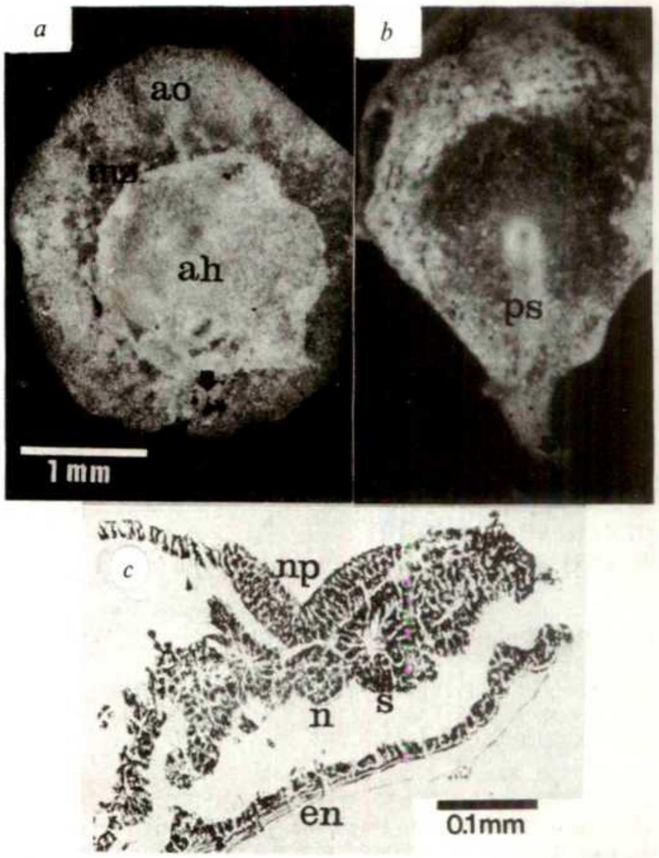


Fig. 1 *a*, The normal hypoblast has been replaced by the aggregated hypoblast (ah) in this stage XIII blastoderm. The marginal zone (mz) and the area opaca (ao) have been left to illustrate the relative area occupied by the aggregate. To prepare aggregated hypoblasts, normal hypoblasts were surgically removed from stage XIII blastoderms. To obtain individual hypoblastic cells, several hypoblasts were incubated in a solution containing trypsin and EDTA, transferred to the incubating medium, dissociated by gentle pipetting so that essentially a single-cell suspension was obtained, centrifuged down and resuspended in the incubating medium at a concentration of about 3×10^5 cells ml^{-1}. The incubating medium consisted of Dulbecco's modified Eagle's minimal essential medium containing 10% heat-inactivated horse serum. The cells were then incubated inside a hollow plastic cylinder whose internal diameter (3 mm) was slightly larger than that of an intact hypoblast. The cylinder of between 1 and 2 cm in height was made to stand on the lower side of a chick egg's vitelline membrane which had previously been stretched on a glass ring and suspended on chick egg albumin. The system was incubated at 37.5 °C in an atmosphere of 95% air/5% CO_2 for between 3 and 11 h. During this period the dissociated hypoblast cells within the cylinder reaggregate at the surface of the membrane and may form a one-layer disk which closely resembles the original hypoblast. The plastic cylinder was then removed, and the reconstituted hypoblast detached from the membrane with fine glass needles and layered on to the ventral side of a normal stage XIII epiblast. *b*, Primitive streak (ps) that developed by layering a reaggregated hypoblast on top of an epiblast proper, previously deprived of the marginal zone and the area opaca. The expected posterior side of the embryo was marked with carbon particles (arrow) on the epiblast before removal of the area opaca. *c*, Transverse section of an embryo that developed from an aggregated hypoblast layered on an epiblast proper. The area opaca and the marginal zone were removed before incubation. np, Neural plate; ch, notochord; s, somite; en, endoderm.

growth of the embryo. The necessity of the interaction with hypoblastic cells but not with the intact polar hypoblast indicates that a normal hypoblast is also acting as an inductor, and consequently that this induction process is indeed a specific function performed by the hypoblast.

The nature of the 'induction' process performed by a normal hypoblast is not yet known. It is unlikely, however, that the hypoblast's function is just one of a physical layer that provides some sort of cover to the denuded epiblast and/or provides non-specific anchorage or guidance for the movement of mesodermal cells. As mentioned above, even a closely associated layer, unless it contains the marginal zone-derived cells, is not sufficient to generate axial structures on a stage XIII chick epiblast.

We have therefore demonstrated that two processes which are independent of each other are necessary for the development of an embryonic axis: a determination of the direction of growth, and a specific induction process which is performed by the hypoblast on the epiblast. Although these two processes are normally both performed by the hypoblast, the second is specific to the hypoblast and can be fully operative even if information concerning the direction of growth is lacking within the hypoblast. We have also demonstrated that in fact there exists a latent information on direction of growth stored in the epiblast which can be expressed by exposing the epiblast to a non-polarized process of induction. The direction of growth information contained in the epiblast is not dominant and is generally overriden by the stronger polarizing effect of a normal hypoblast.

It has also been shown here that a stage XIII epiblast proper can provide sufficient information to organize its own development in a polarized manner. Yet the mesoderm that develops when an epiblast proper is incubated further in the absence of hypoblastic cells, derived from a normal hypoblast, fails to show an axial pattern. It thus seems that only a certain type of mesoderm can be organized in an axial manner and that as a result of the interaction of the epiblast with the hypoblastic cells this particular type of mesoderm emerges.

We thank Y. Shimoni for technical assistance. This work was supported by the Israel Agency for Absorption of Scientists.

Received 17 September; accepted 18 December 1980.

1. Eyal-Giladi, H. & Kochav, S. *Devl Biol.* **49,** 321–337 (1976).
2. Azar, Y. & Eyal-Giladi, H. *J. Embryol. exp. Morph.* **60,** 666–677 (1980).
3. Waddington, C. H. *Phil. Trans. R. Soc.* **B211,** 179–230 (1932).
4. Nicolet, G. *J. Embryol. exp. Morph.* **23,** 79–108 (1970).
5. Spratt, N. T. & Haas, H. *J. exp. Zool.* **144,** 257–276 (1960).
6. Azar, Y. & Eyal-Giladi, H. *J. Embryol. exp. Morph.* **52,** 79–88 (1979).

Generation of cytotoxic T-lymphocyte precursor cells in T-cell colonies grown *in vitro*

Lai-Ming Ching & Richard G. Miller

Department of Medical Biophysics, University of Toronto,
and the Ontario Cancer Institute, 500 Sherbourne Street,
Toronto, Ontario, Canada M4X 1K9

The role of the thymus in T-lymphocyte differentiation remains unclear. The demonstration[1,2] that the thymus can restrict the T-lymphocyte specificity repertoire suggests that T cells acquire specificity within the thymus. However, the demonstrations of immunocompetent helper T cells[3] and cytotoxic T-lymphocyte precursor cells (CLPs)[4] in athymic nude mice suggest that the acquisition of some T-cell reactivity may occur without the thymus. We have been using T-cell colonies grown *in vitro* as a model system for studying various aspects of T-cell differentiation in both mouse[5–8] and man[9]. In one study[6] we showed that CLPs can be found in T-cell colonies grown from spleen cells of normal mice, each colony containing CLPs of several different specificities. The colonies containing CLPs are not clonal, appearing to have a colony-forming unit (CFU-T) of two (perhaps three) cells[6]. Here we provide direct evidence that the CLPs are spontaneously produced in the colonies. In addition, the cells of the CFU-T were characterized with antisera directed against the cell-surface marker Thy-1, which is present on all murine T cells, and the cell-surface markers Lyt-1 and Lyt-2, which are differentially distributed on different T-cell subclasses[10,11]. We found that the CFU-T contains both a Thy-1$^+$ and a Thy-1$^-$ cell, neither of which seems to carry either Lyt-1 or Lyt-2 surface markers.

T-cell colonies were grown from mouse spleen cells as described in Table 1, picked individually after 5–6 days of culture when they had reached a size of 3,000–5,000 cells, and transferred individually into microcultures where they were tested for their ability to generate cytotoxic T lymphocytes (CL) from CLPs. In the experiment shown in Fig. 1, colonies were grown from aliquots of a spleen-cell suspension which either had or had not been depleted of Lyt-2$^+$ cells. As both CL and CLPs are known[10] to carry the surface marker Lyt-2, depletion of Lyt-2$^+$ cells ensured that any CLPs detected in the colonies actually developed there. Both sets of colonies contained CLPs in that, after activation with the mitogen concanavalin A (Con A), they developed cytotoxic activity which could be removed by depletion of Thy-1$^+$ cells. The Thy-1 marker is found on all CL but is absent or only weakly present on a fraction of other types of cytotoxic effector cells[12]. Figure 1*b* shows that when aliquots of the spleen cells used to grow the colonies were tested in a mixed lymphocyte reaction, those depleted of Lyt-2$^+$ cells failed to develop cytotoxic activity, verifying that the depletion treatment did in fact remove CLPs.

Table 1 summarizes experiments designed to test the cellular composition of the CFU-T required for CLP generation. Spleen cells were treated as indicated and then tested for their ability to produce T-cell colonies. If T-cell colonies were obtained, they were individually tested for CLP content as in Fig. 1. Spleen cells, whether depleted of Lyt-1 cells, Lyt-2 cells, both Lyt-1 and Lyt-2 cells, or nylon wool-adherent cells, grew comparable numbers of T-cell colonies of which a comparable fraction contained CLPs. The number of CLPs per colony did not vary appreciably from group to group. Removal of all Thy-1$^+$ cells, however, had a different effect. Although comparable numbers of colonies were obtained, far fewer of them contained CLPs. These colonies were still T-cell colonies in that the cells were killed with anti Thy-1 plus complement (C′). The Thy-1$^+$ cell had to be viable and able to proliferate as the addition of either irradiated or mitomycin C-treated spleen cells to spleen cells depleted of Thy-1$^+$ cells did not increase the CLP content of the colonies. We have shown previously[6] using stimulator cells for activation of specific CLPs in normal spleen colonies rather than Con A activation of all CLPs, that each colony contains CLPs of 10 or more different specificities. These results have been confirmed[13] in experiments in which CLPs were activated using Con A but only CL specific for a particular target were measured.

The above data are consistent with the following model. The central component of the CFU-T is a Thy-1$^-$, Lyt-1$^-$, Lyt-2$^-$ cell which can give rise to Thy-1$^+$ cells in a colony. If the CFU-T also includes a Thy-1$^+$ but Lyt-1$^-$, Lyt-2$^-$ cell, CLPs are also generated in the colony. Although this latter cell seems to have an unusual marker distribution, such a distribution has been previously reported[14]. It is necessary to add the qualification that cells found to lack a marker in a complement-mediated cytotoxicity test may still have this marker but either at a low level or displayed in a manner such that complement-mediated lysis cannot occur[11].

A crucial unresolved question is the role of the thymus in generating the cells required for the CFU-T. Previous studies have shown that colonies containing Thy-1$^+$ cells can be grown

0028-0836/81/090802—03$01.00

Table 1 Cellular composition of CFU-T for CLP generation

Cells plated	Colonies per 5×10^5 cells plated	% Colony cells killed by anti-Thy-1 and C′	Colonies with CLPs / Total colonies %
Spleen	27 ± 10	88	432/564 (77)
Nylon wool non-adherent spleen	68 ± 21	N.D.	84/96 (88)
Lyt-1 depleted spleen	22 ± 8	100	17/24 (71)
Lyt-2 depleted spleen	29 ± 11	85	67/83 (81)
Lyt-1 and Lyt-2 depleted spleen	21 ± 6	N.D.	12/12 (100)
Thy-1 depleted spleen	59 ± 26	98	6/238 (2.5)
Thy-1 depleted spleen plus mitomycin C-treated spleen	42 ± 10	N.D.	1/47 (2.1)
Thy-1 depleted spleen plus irradiated spleen	77 ± 27	N.D.	1/47 (2.1)

Colonies were grown by culturing 5×10^5 viable cells from RNC-*nu*/+ mice treated as indicated, in 1 ml of α-medium[23] supplemented with 0.8% methylcellulose, 40% PHA-LCM, 20% fetal calf serum, 5×10^{-5} M 2-mercaptoethanol and 0.3 mg ml^{-1} L-glutamine as previously described[6]. The PHA-LCM is conditioned medium prepared by incubating human peripheral blood leukocytes (10^6 cells ml^{-1}) for 7 d in α-medium containing 1% PHA and 10% fetal calf serum[16]. Colonies were tested for CLP content as indicated in Fig. 1 except that in some experiments 2×10^5 irradiated (1,500 rads) RNC-*nu*/+ spleen cells were used in place of RNC.*nu* spleen cells and that the cultures were not always supplemented with Con A conditioned medium, in which case peak response was on day 4 and cytotoxic activity per positive colony only half as great. All procedures gave equivalent results with respect to the fraction of colonies containing CLPs[13]. Nylon wool-adherent cells were removed by the procedure of Julius *et al.*[24]. Antiserum treatments on spleen cells were as in Fig. 1. The antisera used, with the addition of hybridoma anti-Lyt-1, clone W3, were also as in Fig. 1. Cells (10^7 per ml) were treated with mitomycin C (25 µg ml^{-1} in medium) by incubation for 20 min at 37 °C. Treatment of colony cells with anti-Thy-1 and C′ was carried out by incubating 10^5 pooled colony cells with anti-Thy-1 (1 : 50 final dilution) and rabbit complement (1 : 15 final dilution) for 45 min at 37 °C. Cell viability was assessed by eosin exclusion.

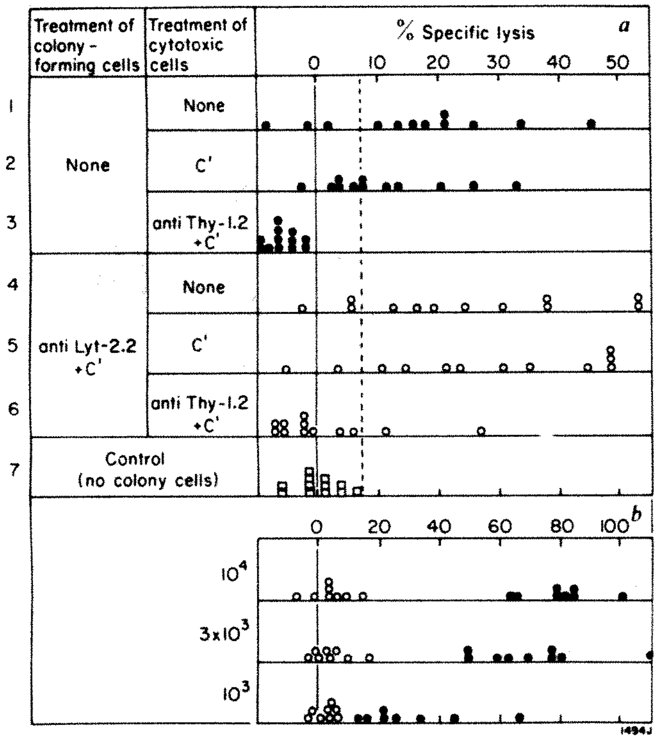


Fig. 1 *a*, Spleen colonies were grown from RNC *nu*/+ spleen cells which either had (○) or had not (●) been treated with anti-Lyt-2 plus complement and were then tested for their ability to produce CL from CLPs by transferring them individually into microwells along with 2×10^5 RNC. *nu* spleen cells in 0.2 ml of α-medium containing 2 µg ml^{-1} Con A and supplemented with 10% fetal calf serum, 5×10^{-5} M 2-mercaptoethanol and 10 mM HEPES buffer as described previously[6]. The mitogen Con A activates all CLPs regardless of specificity[25]. On day 3, 0.1 ml from each culture was replaced with 0.1 ml of conditioned medium from mouse spleen cells stimulated for 18 h with 5 µg ml^{-1} Con A[18] and on day 5, cultures were assayed for cytotoxic activity by adding ^{51}Cr-labelled DBA/2 splenic Con A blasts in the presence of PHA (final dilution 1:60) as described previously[6]. The PHA seems to stick CLs to target cells so that a CL can kill any target cell irrespective of whether it can recognize it[25]. Depletion of Lyt-2 cells was done by incubating 10^7 spleen cells with 50 µl anti-Lyt-2 hybridoma anitserum (clone 19-178) in a total volume of 0.5 ml medium at 4 °C for 30 min, followed by a wash and 45 min incubation at 37 °C with 0.5 ml rabbit complement (1:15 dilution) (Cedarlane Laboratory). Treatment with anti-Thy-1 was done by adding anti-Thy-1 antiserum (1:50 final dilution) and rabbit complement (1:15 final dilution) to the cultures for 45 min at 37 °C. Cultures were washed four times before the addition of ^{51}Cr-labelled target cells. Purified rabbit anti mouse Thy-1 glycoprotein antibody was used[26]. *b*, Three concentrations of the spleen cells used to grow the colonies tested in *a*, both untreated (●) and anti-Lyt-2 plus complement treated (○), were cultured with 3×10^5 irradiated (1,500 rad) DBA/2 spleen stimulators and 10^5 RNC. *nu* spleen cells and assayed on day 5 for cytotoxicity against DBA/2 splenic Con A blasts.

from Thy-1$^-$, presumably pre-thymic, progenitor cells in nude spleen[5], bone marrow[8] and long-term bone marrow cultures[7]. All these colonies seem to contain T-helper cell activity and those of bone marrow origin also contain Thy-1$^+$ cells which are apparently involved in inactivating certain self-reactive CLPs[8]. In common with colonies grown from normal spleen cells depleted of Thy-1$^+$ cells, none of these colonies contains CLPs. We hypothesize that the Thy-1$^+$ component of the CFU-T from normal spleen required to generate CLPs in colonies is a post-thymic cell. It might provide an equivalent to the thymus microenvironment required for CLP generation or itself be a post-thymic CLP precursor cell. According to Stutman[15], immunologically incompetent thymus-processed precursor cells are exported from the thymus to the periphery. The maturation of these post-thymic precursors into competent T cells requires specific interaction with inducer cells in the periphery. Thus, the CFU-T may represent the inducer, post-thymic precursor cell pair.

Growth of the T-cell colonies discussed here has an absolute requirement for a phytohaemagglutinin (PHA)-induced supernatant factor which we call PHA-LCM (see Table 1 legend, refs 16, 17). Although made in much the same manner as IL-2 (also called co-stimulator or T-cell growth factor (TCGF)[18–21]), the incubation period is much longer (7 days compared with 18–48 h). Neither IL-2 nor PHA alone will support growth of T-cell colonies of the type described here[22]. Thus, there seems to be factor(s) in the PHA-LCM which allows the *in vitro* growth and differentiation of more immature T cells than those T cells which can respond in the mixed lymphocyte reaction or can be made to grow in the presence of IL-2. Although the development of functional T cells in the colonies is certainly extra-thymic, this development should probably not be considered as thymus independent, even for those types of colonies for which the CFU-T contains no Thy-1$^+$ cells: the production of PHA-LCM is very likely to depend on the presence of T cells. We conclude that the T-cell colonies described here provide a valuable *in vitro* model system for investigating questions concerning T-cell differentiation and development of the CLP specificity repertoire not approachable with other available *in vitro* techniques.

We thank Dr G. B. Price for discussion and critical comments as well as for providing the PHA-LCM, Dr U. Hämmerling for providing the anti-Lyt-1 and anti-Lyt-2 antisera, and Dr M. Letarte for providing the anti-Thy-1 antiserum. This work was supported by the NCI of Canada.

Received 22 August; accepted 31 December 1980.

1. Zinkernagel, R. M. *et al. J. exp. Med.* **147,** 882–896 (1978).
2. Fink, P. J. & Bevan, M. J. *J. exp. Med.* **148,** 766–778 (1978).
3. Ishikawa, H. & Saito, K. *J. exp. Med.* **151,** 965–968 (1980).
4. Gillis, S., Union, N. A., Baker, P. E. & Smith, K. A. *J. exp. Med.* **149,** 1460–1476 (1979).
5. Jacobs, S. W. & Miller, R. G. *J. Immun.* **122,** 582–584 (1979).
6. Ching, L.-M. & Miller, R. G. *J. Immun.* **124,** 696–701 (1980).

7. Jones-Villeneuve, E. V., Rusthoven, J. J., Miller, R. G. & Phillips, R. A. *J. Immun.* **124,** 597–601 (1980).
8. Muraoka, S. & Miller, R. G. *J. exp. Med.* **152,** 54–71 (1980).
9. Price, G. B., Teh, H.-S. & Miller, R. G. *J. Immun.* **124,** 2352–2355 (1980).
10. Cantor, H. & Boyse, E. A. *J. exp. Med.* **141,** 1376-1389 (1975).
11. Mathieson, B. J., Sharrow, S. O., Campbell, P. S. & Asofsky, R. *Nature* **277,** 478–480 (1979).
12. Herberman, R. B. & Holden, H. T. *Adv. Cancer Res.* **27,** 305–377 (1978).
13. Ching, L.-M. & Miller, R. G. (in preparation).
14. Alter, B. J. & Bach, F. H. *J. Immun.* **123,** 2599–2601 (1979).
15. Stutman, O. *Immun. Rev.* **42,** 138–184 (1978).
16. Lau, L., McCulloch, E. A., Till, J. E. & Price, G. B. *Expl Hemat.* **6,** 114–121 (1978).
17. Price, G. B., Krogsrud, R., Stewart, S. & Senn, J. S. in *Hematopoietic Cell Differentiation* Vol. 10 (eds Golde, D. W., Cline, M. J., Metcalf, D. & Fox, C. F.) 417–431 (Academic, New York, 1978).
18. Gillis, S. & Smith, K. A. *Nature* **268,** 154–156 (1977).
19. Morgan, D. A., Ruscetti, F. W. & Gallo, R. *Science* **193,** 1007–1008 (1976).
20. Paetkau, V., Mills, G., Gerhart, S. & Monticone, V. *J. Immun.* **117,** 1320–1324 (1976).
21. Möller, G. (ed.) *Immun. Rev.* **51,** 1–357 (1980).
22. Ching, L.-M., Muraoka, S. & Miller, R. G. (in preparation).
23. Stanners, C. P., Elicieri, G. & Green, H. *Nature new Biol.* **230,** 52–54 (1971).
24. Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3,** 645–649, (1973).
25. Bevan, M. J., Langman, R. E. & Cohn, M. *Eur. J. Immun.* **6,** 150–156 (1976).
26. Letarte, M. & Meghji, G. *J. Immun.* **121,** 1718–1725 (1978).

Induction of light chain expression in a pre-B cell line by fusion to myeloma cells

Sylvia C. Riley, Emily J. Brock & W. Michael Kuehl

Department of Microbiology, University of Virginia, School of Medicine, Charlottesville, Virginia 22908, USA

Pre-B cells, the first cells in the B-lymphocyte differentiation pathway which express immunoglobulin[1], have recently been shown to express cytoplasmic μ heavy chain (H) but not light chain (L)[2,3]. If, as is believed, pre-B cells are the precursors of immature B lymphocytes, which express surface IgM[4–7], the differentiation of pre-B cells to immature B lymphocytes must be accompanied by the expression of light chains. In this case, it should be possible for the progeny of a single pre-B cell to express a variety of light chains in association with the same heavy chain. We have tested this hypothesis by hybridizing a pre-B cell line 18–81 expressing only cytoplasmic μ chains[8] with variant myeloma cells which do not express light chains. Hybridization of B-lymphoma cells with myeloma cells usually produces a hybrid with the phenotype of the more differentiated parent[9–11]. In this case, the fusion resulted in the induction of light chain expression from the 18–81 genes and we have been able to demonstrate that independent hybrids express different light chains, in accordance with the hypothesis that a pre-B cell committed to expression of a single μ heavy chain can generate progeny expressing different light chains[2,12].

We used a cloned, virus-transformed pre-B-cell line (18–81) which has been shown to express cytoplasmic μ heavy chain (μ_c) but no light chain. Our subline of 18–81 synthesizes up to 1% of its cell protein as μ_c in the absence of detectable light chain synthesis. A hypoxanthine guanine phosphoribosyl transferase (HGPRT)-negative clone of 18–81, 18–81.T5 (μ_c^+, L$^-$), was isolated and hybridized to CBOHC ($\gamma2_a^+$, L$^-$), a thymidine kinase(TK)-negative variant of myeloma CBO($\gamma2_a^+$, κ^+) which, like 18.81.T5, does not express light chain. Immunoglobulin expression was analysed by SDS-polyacrylamide gel electrophoresis of immunoprecipitates isolated from ^{35}S-Met-radiolabelled postnuclear supernatants. Immunoglobulin is expressed by 75% of presumptive hybrids; the reason for lack of immunoglobulin expression by hybrids is unclear but could be due to chromosome segregation. All immunoglobulin-expressing hybrids examined synthesize either or both of the μ and $\gamma2_a$ heavy chains of the parent cells and 13 (40%) of these express a κ light chain which is not produced by either of the parent cells (Fig. 1). Expression of λ light chain is not detected in these hybrids[13]. We tested for λ expression using rabbit antisera specific for mouse λI light chain. The hybrids were not tested for λII or λIII subclasses. To show that expression of κ light chain by these hybrids is not due to re-expression of the wild-type CBO κ light-chain gene, the CBO and hybrid light chains were analysed on an isoelectric focusing gel (Fig. 2). Light chains from three different CBOHC × 18–81.T5 hybrids (Fig. 2, lanes *c–e*) focused at a different *p*H range than the light chain from CBO (Fig. 2*f*). In addition, Southern blots of CBO and CBOHC DNA indicate that the expressed CBO κ light chain gene is deleted from the CBOHC variant (M. Kuehl, unpublished result). Thus expression of κ light chain by these hybrids is not due to re-expression of the wild-type CBO myeloma κ light-chain gene. It must therefore be due to expression of another myeloma κ light-chain gene or a pre-B cell κ light-chain gene. To determine whether the κ light chain expressed in pre-B cell × myeloma hybrids is coded for by the pre-B cell genome, a TK-negative clone of 18–81, 18–81.B4 (μ_c^+, L$^-$), was isolated and hybridized to NT$_2$(H$^-$, L$^-$, F$_{CL}^+$), an HGPRT-negative mutant of myeloma variant NP$_2$ which does not express heavy or light chains[14]. NT$_2$ contains only the aberrantly rearranged κ constant region (C$_\kappa$) gene which is expressed as a 12,000-molecular weight (MW) κ

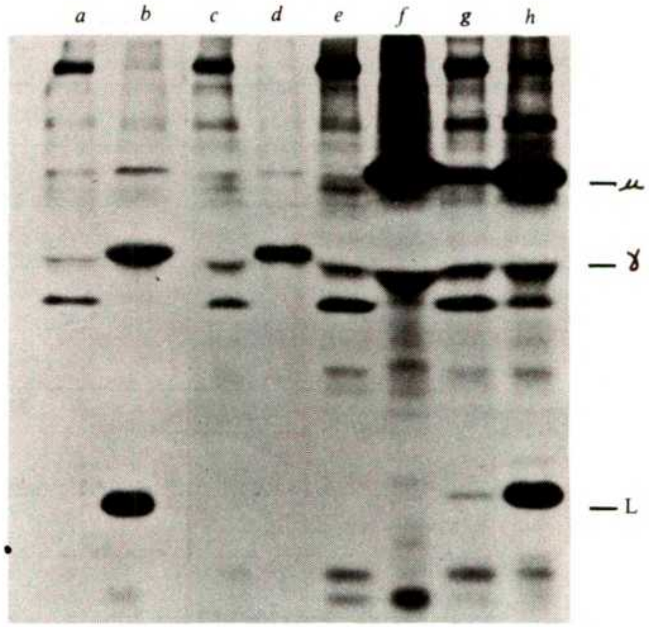


Fig. 1 SDS-polyacrylamide gel electrophoretic analysis of immunoglobulin synthesized by CBOHC myeloma, 18–81.T5 pre-B cell and CBOHC × 18–81.T5 hybrids. The 18–81.T5 HGPRT-negative variant of 18–81 was selected in 6 μg ml^{-1} 2-amino-6-mercaptopurine (6-TG). The CBO ($\gamma2_a^+$, κ^+) TK-negative variant of C1.3 was selected in 30 μg ml^{-1} 5-bromo-2′-deoxyuridine (BUdR) after mutagenesis with 1 μg ml^{-1} *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine. The CBOHC ($\gamma2_a^+$, L$^-$) variant clone was isolated from CBO. 18–81.T5 and CBOHC were fused using 35% polyethelene glycol (PEG) by methods described by Gefter *et al.*[20] and modified by Clafin *et al.*[21]. The parent cells were washed free of serum and mixed in equal numbers (10^7 of each) in 35% PEG. The cells were pelleted twice by low-speed centrifugation, resuspended in growth medium plus HAT (hypoxanthine, aminopterin and thymidine), and cloned into 96-well microtitre plates. The cells and hybrid clones were metabolically labelled for 2 h with ^{35}S-Met, washed and lysed in 1% NP-40 (ref. 22). Postnuclear supernatants were indirectly immunoprecipitated with rabbit antibodies to mouse μ, $\gamma2$, κ and λ chains and goat antiserum to rabbit immunoglobulin[23]. After dissolving the precipitate in SDS the immunoprecipitates were reduced, electrophoresed on 11.25% SDS-Tris-glycine-polyacrylamide gels[24], and analysed by photofluorography[25]. *a, c, e, g,* Normal rabbit serum; *b, d, f, h,* rabbit anti-MPC11 ($\gamma2_b$, κ) + anti-MOPC104 (μ, λ); *a, b,* hybrid 1; *c, d,* CBOHC; *e, f,* 18–81.T5; *g, h,* hybrid 2.

0028–0836/81/090804—03$01.00

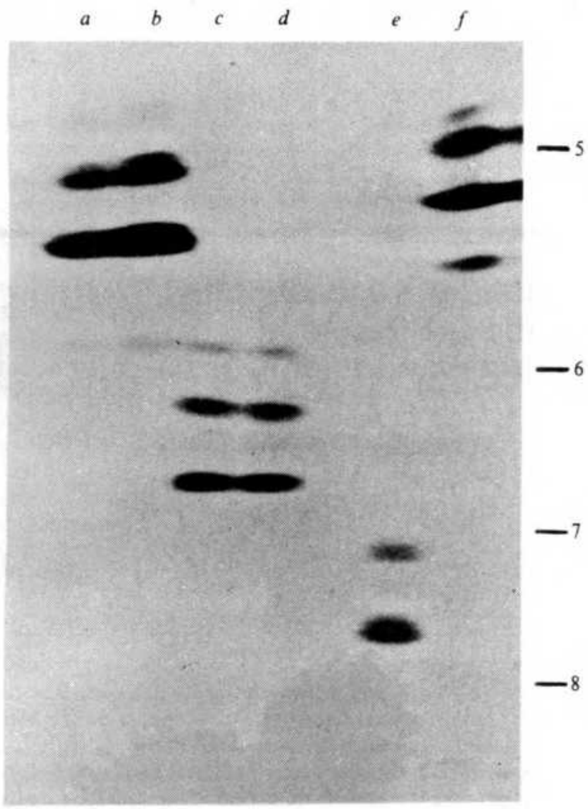


Fig. 2 Isoelectric focusing analysis of light chain from CBO cells and $NT_2 \times$ 18–81.B4 and CBOHC × 18–81.T5 hybrids. Immunoprecipitates using rabbit antisera specific for mouse κ light chain and goat antiserum to rabbit immunoglobulin were prepared from ^{35}S-Met-labelled postnuclear supernatants, reduced and electrophoresed on 9% SDS-polyacrylamide slab gels as described in Fig. 1 legend. A small strip which contained light chain was cut from the gel, equilibrated in 7M urea, 1% NP–40 and 2% ampholines, then focused on an isoelectric focusing gel (4.7% acrylamide, 0.12% *N,N*-methylene bisacrylamide, 16.6% glycerol, 7M urea, 1% NP-40, 2.5% carrier ampholines (*p*H 3.5–10), and 0.033% ammonium persulphate) according to the procedure of Burrows *et al.*[2]. The gel was analysed by photofluorography. *a*, $NT_2 \times$ 18–81. B4 hybrid 1; *b*, $NT_2 \times$ 18–81. B4 hybrid 2; *c*, CBOHC × 18–81.T5 hybrid 3; *d*, CBOHC × 18.81.T5 hybrid 4; *e*, CBOHC × 18–81.T5 hybrid 2; *f*, CBO.

light chain constant region fragment (F_{CL}) in all MPC11 myeloma cells and MPC11 variants[15,16]. Five out of six presumptive hybrids synthesize the μ and F_{CL} chains of the parent cells, but two (40%) of the immunoglobulin-expressing hybrids synthesize a full-sized κ light chain, which migrates slightly more rapidly on an SDS gel than the MPC11 κ light chain (Fig. 3). Expression of λ light chain is not detected in these hybrids[13]. Because the hybrids which express a *de novo* full-sized κ chain also express F_{CL}, and the only C_κ gene present in the NT_2 myeloma parent is expressed as F_{CL}[15,16], the synthesis of a full-sized light chain cannot be due to expression of a myeloma κ light chain gene. Therefore, the κ light chains expressed by the pre-B cell × myeloma hybrids are coded for by the pre-B cell genome.

If the 18–81 pre-B cell is already committed to expression of a particular light chain, then the same light chain should be expressed in all myeloma × pre-B cell hybrids. However, when the κ light chains from five different pre-B cell × myeloma hybrids were compared by isoelectric focusing (Fig. 2) and by peptide mapping of leucine-radiolabelled tryptic–chymotryptic peptides (data not shown), at least three different κ light chains were found. Of the five hybrids examined for the isoelectric pattern of the newly expressed light chain (Fig. 2), there were two pairs of hybrids ($NT_2 \times$ 18–81.B4, Fig. 2*a*, *b*, and CBOHC × 18–81.T5, Fig 2*c*, *d*) with indistinguishable light chains. However, the conclusion that the same light chain is expressed in independent hybrids is not justified because in each case the pair of hybrids expressing apparently identical light chains was from the same hybridization experiment.

Because somatic cell hybrids between myeloma cells and 18–81 express at least three different light chains, induction of light chain expression is probably not due to enhanced expression of a rearranged light-chain gene in the original 18–81 clone. This conclusion is supported by Southern blots, using *Bam*HI-restricted genomic fragments and a nick-translated *Bam*HI-*Xba*I fragment derived from an embryonic C_κ genomic clone as the probe[17]. The C_κ genes in 18–81, 18–81.B4 and 16 subclones of 18–81.B4 are on a 13-kilobase fragment which co-migrates with the 13-kilobase embryonic *Bam*HI fragment present in splenic DNA (data not shown). As DNA rearrangement to form active κ light-chain genes has apparently not occurred in the original 18–81 clone, there are two other possibilities: first, light-chain gene rearrangement could occur spontaneously in the 18–81 population during culture (that is, before somatic-cell hybridization). Second, rearrangement and expression of an 18–81 light-chain gene could also occur during or following hybridization of 18–81 with a myeloma cell. When 24 subclones of 18–81.B4 were analysed for light chain synthesis, none of the subclones expressed κ or λ light chain (D. Pinney and S. Riley, unpublished results). Therefore, if the light chain is rearranged spontaneously in 18–81 cells, fusion of 18–81 with a myeloma cell is required to induce detectable light chain synthesis. This

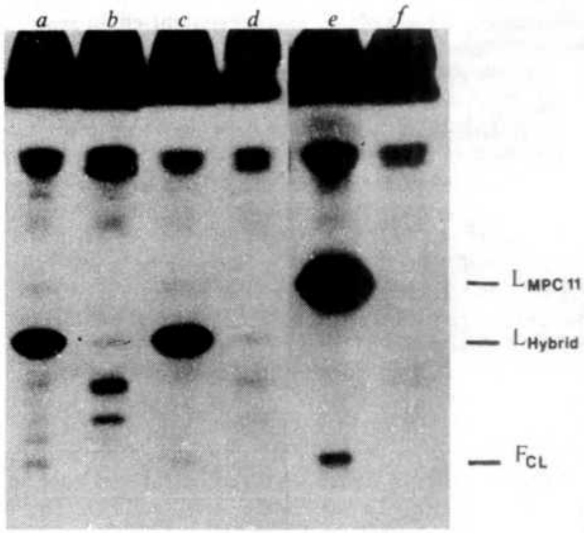


Fig. 3 SDS-polyacrylamide gel electrophoretic analysis of immunoglobulin synthesized by NT_2 myeloma × 18–81.B4 pre-B cell hybrids. The 18–81.B4 TK-negative mutant of 18–81 was selected in 30 μg ml^{-1} BUdR after mutagenesis with 125 μg ml^{-1} ethylmethanesulphonate. 4T00.A(H^-, κ^+, F_{CL}^+) is a spontaneous subline of 4T00 which no longer synthesizes detectable heavy chain. NT_2 (H^-, L^-, F_{CL}^+) is an HGPRT-negative mutant of NP_2. 4T00 (ref. 20) and NP_2 (ref. 14) were derived from the MPC11 myeloma cell line. F_{CL} is expressed in low levels in all MPC11 cell lines and its variants but is not expressed in other myeloma or 18–81 cells (data not shown). Immunoprecipitates were prepared from ^{35}S-Met-labelled postnuclear supernatants, reduced, electrophoresed on a 15% SDS-Tris-glycine-polyacrylamide gel, and analysed by photofluorography as described in Fig. 1 legend. Minor non-immunoglobulin contaminants are seen in the specific and nonspecific immunoprecipitates because the photograph was overexposed to show F_{CL}. *a*, *c*, *e*, Rabbit anti-MPC11 ($\gamma 2_b$, κ) + anti-MOPC104 (μ, λ); *b*, *d*, *f*, normal rabbit serum; *a*, *b*, $NT_2 \times$ 18–81.B4 hybrid 1; *c*, *d*, $NT_2 \times$ 18–81.B4 hybrid 2; *e*, *f*, 4T00.A.

possibility may be supported by the results of Rosenberg *et al.*[18,19] who have observed that lipopolysaccharide (LPS) can induce enhanced μ heavy chain expression and *de novo* light chain expression in some of their 18–81 sublines; although they provide no evidence to show whether LPS is inducing rearrangement of light-chain genes or is only enhancing expression of spontaneously rearranged light-chain genes. In contrast to their results, LPS has no effect on μ heavy chain expression or in inducing light chain expression in our 18–81 subline or the 24 subclones described above (D. Pinney, unpublished results).

Regardless of whether light chain rearrangement occurs spontaneously or as a consequence of cell fusion, the expression of different light chains by different hybrids suggests that when the pre-B cell light chain is expressed in 18–81 × myeloma hybrids, the 18–81 germline C_κ gene can be rearranged to at least three different κ light-chain variable region (V_κ) genes. Southern blots of DNA from pre-B cell × myeloma hybrids provide direct evidence for this rearrangement. One hybrid from each of three different κ light-chain groups, as defined by isoelectric focusing (Fig. 2), was analysed for rearrangement of κ light chains by Southern blots, using *Bam*HI-restricted genomic fragments and a nick-translated *Bam*HI–*Xba*I fragment derived from an embryonic C_κ genomic clone as the probe[17]. Each hybrid lacks the 13-kilobase C_κ fragment but contains a new rearranged C_κ which is not found in either 18–81 pre-B cells or CBOHC, CBO, NT$_2$ or 4T00.A myeloma cells. The *de novo* rearranged C_κ fragments in each of the three hybrids has a unique size compared with the *de novo* C_κ fragments in the other two hybrids (data not shown).

The development of a pre-B cell from a stem cell involves the rearrangement and expression of the μ heavy-chain gene. Rearrangement of a V_L gene to a C_L gene could occur at the same time or after rearrangement of the μ heavy-chain gene. In the former, expression of the rearranged light-chain gene would occur later than expression of the μ heavy-chain gene. In the 18–81 pre-B cell line light-chain genes are not rearranged, although our results indicate that they can be rearranged and expressed. This phenotype may be a result of transformation by Abelson virus, but it probably reflects the phenotype of a pre-B cell population in murine bone marrow, because cells expressing μ heavy chain and no light chain have been found in normal bone marrow[2]. It has been suggested that the asynchronous onset of μ and light-chain synthesis could be involved in the expression of antibody diversity[2]. That is, a pre-B cell selects a particular V_H gene and rearranges it next to the C_μ gene. The progeny of this pre-B cell may then select a V_L gene to translocate it next to a C_L gene. In this way a rapidly proliferating pre-B cell which expresses a specific heavy-chain V region is the precursor of cells which express different light chains along with this heavy chain. Our results strongly support this hypothesis.

We thank Mr James Remenick for technical assistance and N. Rosenberg and D. Baltimore for the 18–81 cell line. This work was supported by NIH research grant AI12525. S.C.R. is supported by NIH training grant 2T32-GM07082. W.M.K. is a recipient of a Research Career Development Award AI00293.

Received 15 August; accepted 18 December 1980.

1. Raff, M. D., Megson, M., Owen, J. J. T. & Cooper, M. D. *Nature* **259,** 224–226 (1976).
2. Burrows, P., LeJeune, M. & Kearney, J. F. *Nature* **280,** 838–841 (1979).
3. Levitt, D. & Cooper, M. D. *Cell* **19,** 617–625 (1980).
4. Osmond, D. G. & Nossal, G. J. V. *Cell. Immun.* **13,** 132–145 (1974).
5. Ryser, J. E. & Vassalli, P. *J. Immun.* **113,** 719–728 (1974).
6. Owen, J. J. T., Wright, D. E., Habu, S., Raff, M. C. & Cooper, M. D. *J. Immun.* **118,** 2067–2072 (1977).
7. Burrows, P. D., Kearney, J. F., Lawton, A. R. & Cooper, M. D. *J. Immun.* **120,** 1526–1531 (1978).
8. Siden, E. J., Baltimore, D., Clark, D. & Rosenberg, N. E. *Cell* **16,** 389–396 (1979).
9. Eshhar, Z., Blatt, C., Bergman, Y. & Haimovich, J. *J. Immun.* **122,** 2430–2434 (1979).
10. Laskov, R., Kim, K. J. & Asofsky, R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 915–919 (1979).
11. Raschke, W. C., Mather, E. L. & Koshland, M. E. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3469–3473 (1979).
12. Cooper, M. D., Burrows, P. D. & Lawton, A. R. in *Biological Basis of Immunodeficiency* (eds Gelfand, E. W. & Dosch, H. M.) 57–69 (Raven, New York, 1980).
13. Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209,** 1366–1369 (1980).
14. Coffino, P. & Scharff, M. D. *Proc. natn. Acad. Sci. U.S.A.* **68,** 219–223 (1971).
15. Choi, E., Kuehl, M. & Wall, R. *Nature* **286,** 776–779 (1980).
16. Seidman, J. G. & Leder, P. *Nature* **286,** 779–783 (1980).
17. Seidman, J. G. & Leder, P. *Nature* **276,** 790–795 (1978).
18. Baltimore, D., Rosenberg, N. & Witte, O. N. *Immun. Rev.* **48,** 3–22 (1979).
19. Rosenberg, N., Siden, E. & Baltimore, D. in *B Lymphocytes in the Immune Response* (eds Cooper, M., Mosier, D. E., Scher, I. & Vitetta, E. S.) 379–386 (Elsevier, New York, 1979).
20. Gefter, M. L., Margulies, D. H. & Scharff, M. D. *Somat. Cell Genet.* **3,** 231–236 (1977).
21. Clafin, L. & Williams, K. *Curr. Topics Microbiol. Immun.* **81,** 107–109 (1978).
22. Kuehl, W. M. & Scharff, M. D. *J. molec. Biol.* **89,** 409–421 (1974).
23. Horowitz, M. & Scharff, M. D. in *Fundamental Techniques in Virology* (eds Habel, K. & Salzman, N. P.) 235–315 (Academic, New York, 1969).
24. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
25. Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56,** 335–341 (1975).

Reaction of 5-iodonaphthyl-1-nitrene with the IgE receptor on normal and tumour mast cells

David Holowka*, Carlos Gitler†, Tuvia Bercovici† & Henry Metzger*

* Section on Chemical Immunology, Arthritis and Rheumatism Branch, NIAMDD, NIH, Bethesda, Maryland 20205, USA
† Department of Membrane Research, The Weizmann Institute of Science, Rehovot, Israel

Mast cells, basophils and a tumour analogue—rat basophilic leukaemia (RBL) cells—have a surface glycoprotein (R_ε) which specifically binds monomeric immunoglobulin E (IgE)[1], and aggregation of R_ε causes secretion[1–3]. When isolated from non-ionic detergent extracts of surface-labelled RBL cells by IgE-specific immunoprecipitation, R_ε appears as a 50,000 (50 K) to 60,000 (60 K) molecular weight (MW) band on electrophoresis in polyacrylamide gels in SDS (SDS-PAGE)[4,5]. Likewise, only a 50 K component is observed when the polypeptide that binds IgE is isolated by affinity chromatography in conditions which prevent aggregation of the IgE[6–8], even when intrinsically labelled R_ε is studied[9]. To determine how R_ε is inserted into the plasma membrane, we reacted RBL cells with the photolysable hydrophobic reagent 5-iodonaphthyl-1-azide (INA), which preferentially labels the intramembranous segments of several intrinsic membrane proteins[10,11]. We report here that, surprisingly, the label was found, not on the 50 K glycopeptide, but only on a 30 K component which other studies suggest is a subunit of R_ε (refs 9, 12).

RBL cells from the 2H3 subline were grown as described previously[13]. In some experiments cells were first incubated with IgE. Before reaction with INA, the cells were washed to remove albumin present in the serum-containing incubation medium, with which the INA can react[10,11]. Normal rat mast cells were saturated with IgE and isolated from peritoneal washings of Sprague–Dawley rats as described by Holgate *et al.*[14], omitting the final gradient step.

Radioiodinated INA was prepared on the day of the experiment from 5-amino-1-azidonaphthalene[10]. Solutions containing INA were irradiated at 25–30 °C in a water-jacketed 1-cm Pyrex cuvette with a HB0 200W 14 Osram mercury lamp at a distance of 10 cm. The light was focused with a fused quartz lens and passed through Corning CS7-60 and 054 filters to yield a band centred at ~350 nm with a halfwidth of 60 nm. In all the experiments 10^7 cells in phosphate-buffered saline, *p*H 7.4, were irradiated for 8 min in five 2-ml batches in 3 μm INA (specific activity ~1 Ci mmol^{-1}), 0.02% gelatin and 10 mM neutralized glutathione. The cells were then diluted twofold in medium containing 20% fetal calf serum and were ≥80% viable as judged by dye exclusion. The cells were washed twice and then solubilized with non-ionic detergent. IgE was added to the extracts from those cells which had not been previously reacted with it.

Figure 1 shows an experiment in which ^{125}I-INA was photolysed in the presence of RBL cells preloaded with IgE. The major band of ^{125}I counts had an apparent MW of 30 K. No discrete ^{125}I peaks occurred where a band from surface-labelled cells appears (50 K), or where the ^{131}I label on receptor-bound IgE appeared in this gel—at 80 K and 24 K. Extracts from cells incubated with IgE after reaction with INA gave indistinguishable results. Control (human IgE) precipitates contained only very low molecular weight labelled material which may represent phospholipids[11] or nonspecific small peptides.

The 50 K glycopeptide component of R_ϵ is associated specifically and stoichiometrically with a 30 K polypeptide after intact RBL cells or detergent extracts of the cells are cross-linked with dimethyl-3,3′-dithiobispropionimidate (DTBP)[12]. We tested whether the 30 K component labelled with INA might represent the same polypeptide. Cells were incubated with ^{125}I-labelled IgE amidinated to prevent its cross-linking to the receptor by DTBP at a later step[12]. The washed cells were reacted with ^{131}I-labelled INA, the detergent extract reacted with DTBP and immunoprecipitates analysed both before and after reduction with 0.25 M 2-mercaptoethanol (Fig. 2). The reduced sample showed label at an apparent MW of 30 K but none where surface-labelled receptor appears on such gels (~55 K). The unreduced sample showed only a single band for those ^{131}I counts which penetrated the stacking gel, having an apparent MW of ~90 K, exactly where cross-linked R_ϵ from surface-labelled cells appears[12]. In matched non-cross-linked samples label was observed in a 30 K but not in a 50–60 K component.

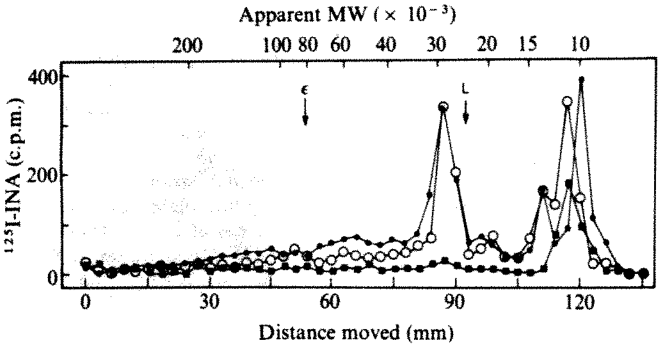


Fig. 1 SDS-PAGE patterns of immunoprecipitates from Triton X-100 extracts of cells labelled with ^{125}I-INA. Cells were lysed with 0.5% detergent in borate-buffered saline (*p*H 8.5) containing 0.5 mM phenylmethylsulphonyl fluoride and 0.5 units ml^{-1} aprotinin. After centrifugation at 50,000*g* for 30 min the extracts were diluted to 0.25% detergent and reacted with human IgE and anti-human IgE plus a *Staphylococcus aureus* Cowan preparation. The supernatant from this immunoprecipitation was then reacted with anti-rat IgE plus *S. aureus* Cowan in a similar manner. The precipitates were washed, eluted with 2% SDS and then reduced in 0.25 M 2-mercaptoethanol. ■, Extract from cells reacted with IgE before reaction with INA and precipitated with anti-human IgE and human IgE as a control. ●, Supernatant from precipitation in (■), precipitated with anti-rat IgE. ○, Extract from cells reacted with INA before addition of IgE and precipitated with anti-rat IgE (after a human IgE–anti-human IgE 'clearing' precipitation). The arrows at 80 K and 24 K are the positions at which the peaks of the ^{125}I-labelled ε and light chains of the rat IgE appeared. Electrophoresis was carried out on a 2.5–27% Isophore gradient gel (Isolab) as described elsewhere[17].

As 1 mol of R_ϵ binds 1 mol of IgE[1,8], we could compare the recovery of ^{131}I-INA-labelled components relative to the ^{125}I-IgE in the same samples. Equivalent amounts of the 30 K component were found for the reduced samples, both cross-linked and non-cross-linked. A lower amount was seen in the 90 K band of the cross-linked unreduced sample and is probably accounted for by incomplete penetration of the cross-linked receptor into the gel, which has occasionally been observed[12]. However, most of the counts which failed to penetrate the stacking gel are not due to cross-linked receptor; such material was seen variably, when it was observed it was always present in the control precipitates also, and was generally largely absent after reduction.

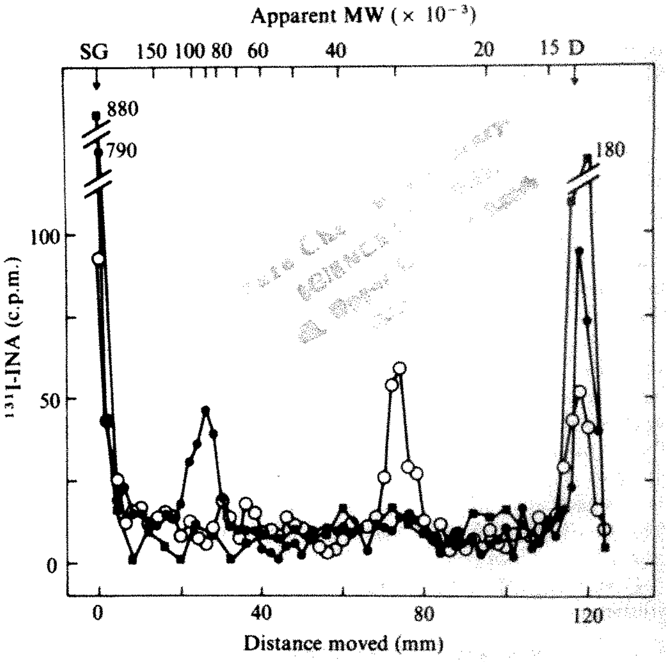


Fig. 2 SDS-PAGE patterns of immunoprecipitates from a Triton X-100 extract of RBL cells reacted with ^{131}I-INA after loading with amidinated ^{125}I-IgE. The supernatant from a 50,000*g* centrifugation was reacted with DTBP in two doses over 4 h, giving a final concentration of 10 mM. The cross-linking reagent was quenched with 100 mM ammonium acetate and the extract was then immunoprecipitated as described in Fig. 1 legend. ■, Control precipitate; ●, ○, anti-rat IgE precipitates. ○, Reduced sample; ●, ■, unreduced samples. The samples were electrophoresed on a 12% polyacrylamide gel as described elsewhere[12]. SG, stacking gel; D, bromophenol blue dye front.

Labelling of R_ϵ on normal rat mast cells gave comparable results (Fig. 3). Two peaks were observed in the specific precipitate only—a major peak at 32 K and a smaller one at 23 K. We have evidence[9] that the latter is a proteolytic fragment of the 30 K component.

R_ϵ behaves like an integral membrane protein: it requires non-ionic detergents to be solubilized and its partial specific volume suggests that it binds detergent (or lipids)[15]. Its resistance *in situ* to proteolysis[16] and its failure to be surface-labelled

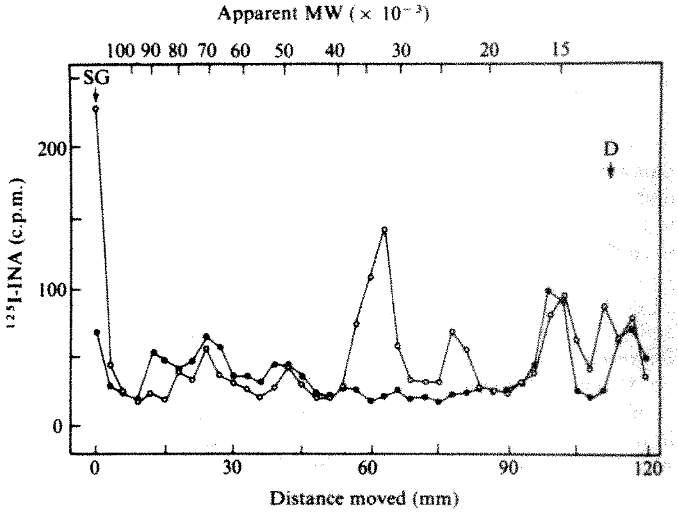


Fig. 3 SDS-PAGE patterns of INA-labelled R_ϵ from normal rat mast cells (~90% purity) preincubated with rat IgE. A Triton X-100 extract of the cells was immunoprecipitated as in Figs 1 and 2. ●, Human IgE–anti-human IgE 'clearing' precipitate; ○, specific anti-rat IgE precipitate. The SDS eluates were reduced before electrophoresis on a 12.5% polyacrylamide gel.

when IgE is bound to it[4] are also consistent with a substantial proportion of the receptor being within the plasma membrane or on its cytoplasmic side. The failure of the 50 K subunit to become labelled may result from its having only a short intramembranous segment; this segment may be interacting tightly with the 30 K subunit or membrane lipids, or its amino acid residues may be relatively unreactive. Other studies suggest preferential labelling of unsaturated over saturated lipids by INA[11] and preferential labelling of certain amino acid side chains is possible. Alternatively, the 50 K component may be anchored to the membrane solely through its interaction with the 30 K component.

Cross-linking studies had suggested that R_ε might have two subunits[12]. However, cross-linking reagents may fail to reveal associations because they depend on the apposition of suitably reactive residues. The present study shows that labelling with INA and analogous compounds may provide an alternative method for revealing presumptive subunits of membrane proteins which are undetectable by surface labelling and which may be lost during the rigorous purification often required when intrinsic labelling is used.

D.H. was supported by the Arthritis Foundation during these studies.

Received 15 September; accepted 16 December 1980.

1. Metzger, H. *Immun. Rev.* **41,** 186–199 (1978).
2. Ishizaka, T., Chang, T. H., Taggart, M. & Ishizaka, K. *J. Immun.* **119,** 1589–1596 (1977).
3. Isersky, C., Taurog, J. D., Poy, G. & Metzger, H. *J. Immun.* **121,** 549–558 (1978).
4. Conrad, D. H. & Froese, A. *Immun.* **116,** 319–326 (1976).
5. Kulczycki, A. Jr, McNearney, T. A. & Parker, C. W. *J. Immun.* **117,** 661–665 (1976).
6. Conrad, D. H. & Froese, A. *Immunochemistry* **15,** 283–288 (1978).
7. Kanellopoulos, J., Rossi, G. & Metzger, H. *J. biol. Chem.* **254,** 7691–7697 (1979).
8. Kanellopoulos, J. M., Liu, T. Y., Poy, G. & Metzger, H. *J. biol. Chem.* **255,** 9060–9066 (1980).
9. Holowka, D. & Metzger, H. (submitted).
10. Bercovici, T. & Gitler, C. *Biochemistry* **17,** 1484–1489 (1978).
11. Bercovici, T. & Gitler, C. *Ann. N.Y. Acad. Sci.* (in the press).
12. Holowka, D., Hartmann, H., Kanellopoulos, J. & Metzger, H. *J. Receptor Res.* **1,** 41–68 (1980).
13. Taurog, J. D., Fewtrell, C. & Becker, E. *J. Immun.* **122,** 2150–2153 (1979).
14. Holgate, S. T., Lewis, R. A. & Austen, K. F. *J. Immun.* **124,** 2093–2099 (1980).
15. Newman, S. A., Rossi, G. & Metzger, H. *Proc. natn. Acad. Sci. U.S.A.* **74,** 869–872 (1977).
16. Metzger, H., Budman, D. & Lucky, P. *Immunochemistry* **13,** 417–423 (1976).
17. Plotz, P. H., Kimberly, R. P., Guyer, R. L. & Segal, D. M. *Molec. Immun.* **16,** 721–729 (1979).

Expression of the *E. coli uvrA* gene is inducible

Cynthia J. Kenyon & Graham C. Walker

Biology Department, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

***UvrA*⁺-dependent excision repair[1,2] is one of the most important systems in *Escherichia coli* for repairing UV-induced pyrimidine dimers[1] and a variety of other forms of DNA damage[3,4]. The *uvrA* protein acts in conjunction with the *uvrB* and *uvrC* gene products to introduce a nick at the site of a DNA lesion and thus initiate the repair process[1]. We have recently used the Mud(Ap, *lac*)[5] operon fusion vector to identify a set of genes whose expression is induced by DNA damage. One Mud(Ap, *lac*) insertion mapped at the *uvrA* locus and made the cells sensitive to UV light. In this fusion strain, β-galactosidase expression was induced by DNA-damaging agents in a *recA*⁺*lexA*⁺-dependent fashion[6]. We were surprised by this result because *uvrA*⁺-dependent excision repair is observed both in cells in which protein synthesis has been inhibited[7] and in *recA*⁻ and *lexA*⁻ cells[8], findings which have led to the conclusion that the *uvrA* gene product is constitutively expressed and not under the control of the complex *recA*⁺*lexA*⁺ regulatory circuitry[9] (see below). We have investigated this possibility further and describe here the generation and characterization of a set of fusions of the *lac* genes to the promoter of the *uvrA* gene. We confirm that the *uvrA* gene product is induced by DNA damage in a *recA*⁺*lexA*⁺-dependent fashion.**

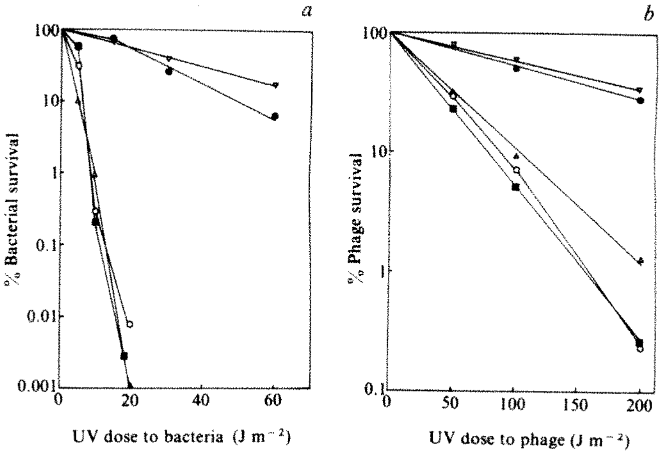


Fig. 1 *a*, Survival after UV irradiation. Fusion strains were grown in supplemented M9 minimal media[11] to a density of 1×10^8 cells ml^{-1}. Aliquots were UV-irradiated as indicated. Survival was determined by plating appropriate dilutions on supplemented M9 plates. *b*, Host cell reactivation. λcI857 was irradiated in 10 mM $MgSO_4$. 0.1 ml diluted phage was preadsorbed to 0.1 ml of an exponential culture of the indicated strain. Infective centres were scored after plating on tryptone plates and incubating at 30 °C. ▽, GW1000 (ref. 6) (parent of fusion strains); △, AB1886 (ref. 12) (*uvrA6*); ○, *uvrA215*::Mud(Ap, *lac*); ●, *uvrA215*::Mud(Ap, *lac*)/pDR2000; ■, *uvrA215*::Mud(Ap, *lac*)/pDR2026.

We were concerned that as a result of some secondary event, such as a Mu-mediated deletion[10], the Mud(Ap, *lac*) insertion at the *uvrA* locus could have become fused to the regulatory region of some gene other than *uvrA*. We therefore decided to obtain Mud(Ap, *lac*) insertions in the *uvrA* gene without imposing the requirement that the resulting fusions possess specific regulatory characteristics. Thus, we made a search for *uvrA*::Mud(Ap, *lac*) insertion mutants based on the criterion of UV sensitivity. Colonies containing random insertions of Mud(Ap, *lac*) were replica plated and UV irradiated (10 J m⁻²). Among 110,000 colonies screened, 55 UV sensitive mutants were obtained. Of these, 13 mapped at the *uvrA* locus (~50% linked to *malB* by bacteriophage P1 transduction[11]). The UV sensitivity of one of these putative *uvrA*::Mud(Ap, *lac*) isolates is shown in Fig. 1.

Like known excision repair mutants, they showed (1) a marked decrease in their ability to support the growth of UV-irradiated bacteriophage such as λ (host-cell reactivation[12]) (Fig. 1), (2) hypermutability with UV light[13] (data not shown), (3) wild-type levels of survival and mutagenesis on exposure to the methylating agent methyl methanesulphonate[14] (unpublished results).

To demonstrate conclusively that these isolates were indeed *uvrA* mutants, we carried out complementation tests. The fusion strains were transformed with the plasmids pDR2000 and pDR2026, which are derivatives of pBR322 carrying the *uvrA*⁺ gene and a *uvrA*::γδ mutation, respectively[15]. As shown in Fig. 1, pDR2000 (but not pDR2026) restored wild-type UV resistance and host-cell reactivation levels. Thus, these isolates carry Mud(Ap, *lac*)-generated insertion mutations of *uvrA*.

The original *uvrA*::Mud(Ap, *lac*) insertion mutant displayed a substantial level of β-galactosidase. The 13 new *uvrA*::Mud(Ap, *lac*) mutants fell into two classes: eight displayed levels similar to the original strain and five showed very low β-galactosidase levels. The two classes were shown to contain Mud(Ap, *lac*) insertions in opposite orientations by chromosomal mobilization experiments[11] (Table 1).

The two classes of *uvrA*::Mud(Ap, *lac*) mutants were tested for mitomycin C inducibility of β-galactosidase. As expected, β-galactosidase levels in strains in which *lac* was integrated opposite to the direction of transcription were unaffected by mitomycin C. However, the eight isolates containing *lac* oriented in the direction of transcription each showed a stimulation of activity on exposure to the chemical (Fig. 2*a*). These same strains showed a similar inductive response to UV irradiation and treatment with nalidixic acid. Thus, transcription of the *uvrA* gene seems to be stimulated by agents that either damage DNA or block DNA replication.

The induction of *uvrA* is only one of several bacterial responses to DNA damage. Treatment of cells with a variety of DNA-damaging agents (including those mentioned above) triggers expression of the 'SOS' functions. These include the induction of certain prophages such as λ, the induction of the *recA* protein, and filamentous growth[16]; at least some of these 'SOS' responses may be due to increased expression of the set of *din* genes[6]. All these phenomena are coordinately regulated by the *recA* and *lexA* proteins. The *lexA* protein has been identified as the repressor of the *recA* gene[17] (D. W. Mount and J. W. Little; R. Brent and M. Ptashne, personal communication); the *lexA* gene product has a negative regulatory effect on the *lexA* gene itself[18] as well as on *din* genes (C.J.K. and G.C.W., unpublished). The regulatory effects of *recA* seem to be due to the proteolytic activity of the *recA* protein which can cleave both the *lexA* protein[19] and λ repressor[20].

We have examined the mechanism by which *recA* and *lexA* act to regulate the *uvrA* gene product. Induction is controlled by the 'SOS'-regulatory circuitry, as the introduction of *recA*$^-$ and *lexA*$^-$ (uninducible repressor) mutations prevents increased β-galactosidase expression in the fusion strains (Fig. 2*a*). Furthermore, an allele of *recA* (known as *lexB30*, ref. 21), which specifies a *recA* protein deficient in proteolytic activity but not recombination functions (J. Roberts, personal communication), also reduces the induction (data not shown). This suggests that the *recA* protein is functioning as a protease in the induction process.

If, like *recA* and *lexA*, *uvrA* is directly repressed by the *lexA* protein, then *uvrA*::Mud(Ap, *lac*) fusion strains lacking the *lexA* gene product should produce high levels of β-galactosidase regardless of whether the *recA* protein is proteolytically active. As shown in Fig. 2*b*, cells carrying a putative null allele of *lexA* called *spr* (ref. 17) synthesize high levels of β-galactosidase in a *recA*$^-$ background. Thus, the *lexA* protein may also repress the *uvrA* gene, with induction occurring when the *recA* protein proteolytically inactivates the *lexA* protein in response to DNA damage.

Table 1 Orientation of Mud(Ap, *lac*) in representative *uvrA*::Mud(Ap, *lac*) insertion mutants

	Pur$^+$ recombination per 10^6 donor cells	
	Donor	
Recipient	*uvrA215*:: Mud(Ap, *lac*)/ $F_{ts1141lac}$	*uvrA223*:: Mud(Ap, *lac*)/ $F_{ts1141lac}$
purD::Mud(Ap, *lac*)	510	48
purA::Mud(Ap, *lac*)	2	1,780

$F_{ts114lac}$ (ref. 27) was introduced into GW1000 derivatives carrying *uvrA215*::Mud(Ap, *lac*) (Lac$^+$, inducible) and *uvrA233*:: Mud(Ap, *lac*) (Lac$^-$, uninducible). Representative members of the two classes would be expected to transfer host genes in opposite directions when an F' *lac* was integrated into the *lac* gene of Mud(Ap, *lac*) to form an Hfr. Cultures of these strains contain a small fraction of Hfrs resulting from F' integration into Mud(Ap, *lac*) via *lac* homology. The resulting gradient of transfer of chromosomal markers depends on the orientation of Mud(Ap, *lac*) within *uvrA*. Recipients were derivatives of BW545 (*lacΔU169 rpsL1*) bearing *purA*::Mud(Ap, *lac*) and *purD*::Mud(Ap, *lac*) insertions (E. Trefelner and G.C.W. unpublished). (The *purA* and *purD* genes map on either side of *uvrA*.) The presence of Mud(Ap, *lac*) in the recipient prevented killing due to zygotic induction of Mu functions. Equal volumes of exponential cultures were mixed and incubated at 34 °C for 30 min. Matings were interrupted by vortexing and 0.1-ml aliquots were plated on minimal plates lacking adenine and donor nutrients. Plate matings were prevented by shifting the cells to 43 °C after 3 h at 34 °C; only Pur$^+$ recombinants escape killing by Mud(Ap, *lac*) *cts* induction. Given the known orientation of *lac* on $F_{ts114lac}$, we conclude that *uvrA* is transcribed in an anticlockwise direction on the *E. coli* chromosome.

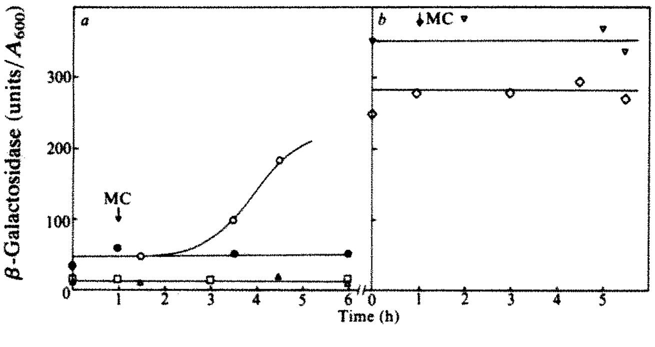


Fig. 2 Kinetics of induction of β-galactosidase in *uvrA*::Mud(Ap, *lac*) fusion strains. Exponential cultures growing in supplemented M9 media were treated with MC (1 μg ml^{-1}) as shown and the specific activity of β-galactosidase determined as described elsewhere[6]. *a*, ●, *uvrA215*::Mud(Ap, *lac*), no MC; ○, *uvrA215*::Mud(Ap, *lac*); □, *recA uvrA215*::Mud(Ap, *lac*); △, *lexA uvrA215*::*Mud*(*Ap, lac*); *b*, ▽, *spr uvrA215*::Mud(Ap, *lac*); ◇, *spr recA uvrA215*::Mud(Ap, *lac*). The effect of the *uvrA* gene product on induction was determined by introducing pDR2000 into wild-type, *lexA*$^-$ and *recA*$^-$ cells. No significant differences in induction profiles were observed (data not shown).

Our suggestions concerning the inducibility of *uvrA* have been recently directly confirmed by Kacinski, Rupp and Seeberg (personal communication), who have found that *uvrA* protein is constitutively expressed at high levels in a *spr recA* strain, and that synthesis of *uvrA* protein is induced by nalidixic acid, a treatment which induces SOS functions.

During our search for UV-sensitive mutants, we isolated two putative *uvrB*::Mud(Ap, *lac*) fusions. These (1) map at the *uvrB* locus (20% co-transducible with *gal* by bacteriophage P1 transduction), (2) give rise to the same set of excision repair-deficient phenotypes as do the *uvrA* insertions, and (3) synthesize β-galactosidase. Although, curiously, these fusions show little mitomycin C inducibility, they display levels of UV and nalidixic acid inducibility comparable with the *uvrA* fusion strains (data not shown).

The findings that components of the excision repair system are synthesized at increased levels in response to DNA damage imply that *E. coli* induces certain repair capabilities as part of its overall defence strategy. The inducibility of excision repair components helps explain certain earlier observations. First, inactivation of the *lexA* protein (which results in the induction of the *uvrA* gene product) increases the UV resistance of *recA* mutants; this increased resistance is dependent on a *uvrA*$^+$ genotype[22]. Second, *recF* and *umuC* mutants show decreased levels of inducible (Weigle[16]) reactivation of UV-irradiated bacteriophage; in each case a *uvrA* mutation eliminates the remaining activity[23,24].

The specific mechanism by which induction of the *uvrA* gene product enhances cell survival is unclear. In conditions in which *uvrA*$^+$-dependent inducible UV resistance is observed, there seems to be no increase in the levels of 'short patch' excision repair[25]. Thus, these proteins may be functioning in another repair process. One such possibility is another type of excision repair called 'long patch' repair[26]. Like the induction of the *uvrA* gene product, long patch repair is *recA*$^+$*lexA*$^+$ dependent.

We thank A. Bagg, S. Elledge, P. Langer, J. Mulligan, P. Pang, K. Perry, W. Shanabruch, E. Trefelner and S. Winans for valuable discussions, L. Withers for help in preparing the manuscript, and especially B. M. Kacinski, W. D. Rupp, E. Seeberg, P. K. Cooper, J. Roberts, D. M. Mount and J. W. Little for sharing information before publication. This work was supported by a Rita Allen Scholar Award to G.C.W. and an NSF predoctoral fellowship to C.J.K.

Received 19 August; accepted 5 December 1980.

1. Seeberg, E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2569–2573 (1978).
2. Hanawalt, P. C., Cooper, P. K., Ganesan, A. K. & Smith, C. A. *A. Rev. Biochem.* **48,** 783–836 (1979).
3. Ikenaga, M., Ichikawa-Ryo, H. & Kondo, S. *J. molec. Biol.* **75,** 341–356 (1975).
4. Boyce, R. P. & Howard-Flanders, P. *Z. VerebungLehre.* **95,** 345–350 (1969).
5. Casadaban, M. J. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4530–4533 (1979).
6. Kenyon, C. J. & Walker, G. C. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2819–2823 (1980).
7. Boyle, J. M. & Setlow, R. B. *J. molec. Biol.* **51,** 131–144 (1970).
8. Clark, A. J. & Volkert, M. R. in *DNA Repair Mechanisms* (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 57–72 (Academic, New York, 1978).
9. Radman, M. in *Molecular Mechanisms for Repair of DNA* (eds Hanawalt, P. C. & Setlow, R. B.) 355–367 (Plenum, New York, 1975).
10. Faelen, M., Huisman, O. & Toussaint, A. *Nature* **271,** 580–582 (1978).
11. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
12. Howard-Flanders, P., Boyce, R. P. & Theriot, L. *Genetics* **53,** 1137–1150 (1966).
13. Ishii, Y. & Kondo, S. *Mutat. Res.* **27,** 27–44 (1975).
14. Kondo, S., Ichikawa, H., Iwo, K. & Kato, T. *Genetics* **66,** 187–217 (1970).
15. Sancar, A. *et al. J. molec. Biol.* (submitted).
16. Witkin, E. M. *Bact. Rev.* **40,** 869–907 (1976).
17. Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 300–304 (1977).
18. Brent, R. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1932–1936 (1980).
19. Little, J. W., Edmiston, S. H., Pacelli, L. Z. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3225–3229 (1980).
20. Roberts, J., Roberts, C. & Craig, N. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4714–4718 (1978).
21. Morand, P., Blanco, M. & Devoret, R. *J. Bact.* **131,** 572–582 (1977).
22. Mount, D. W., Walker, A. C. & Kosel, C. *J. Bact.* **121,** 1203–1207 (1975).
23. Rothman, R. H., Margossian, L. J. & Clark, A. J. *Molec. gen. Genet.* **169,** 279–287 (1979).
24. Walker, G. C. & Dobson, P. P. *Molec. gen. Genet.* **172,** 17–24 (1979).
25. Swenson, P. A. & Setlow, R. B. *J. molec. Biol.* **15,** 201–219 (1966).
26. Cooper, P. K. & Hanawalt, P. C. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1156–1160 (1972).
27. Low, K. B. *Bact. Rev.* **36,** 587–607 (1973).

Binding of an SV40 T antigen-related protein to the DNA of SV40 regulatory mutants

R. McKay

Cold Spring Harbor Laboratory, Cold Spring Harbor,
New York 11724, USA

D. DiMaio

Department of Microbiology, Johns Hopkins University School of
Medicine, Baltimore, Maryland 21205, USA

When simian virus 40 (SV40) infects host cells, the viral genes are expressed in two phases. Early in infection, the virus genome is transcribed and translated to yield two related proteins, 'large T' and 'small t' antigens. Later the rate of synthesis of the early messenger RNA falls and the viral genes coding for the coat proteins are transcribed and translated. Early gene transcription is inhibited only when an active large T antigen is present in the cell (for review see ref. 1). These and other observations have led to the hypothesis that T antigen acts as a negative feedback repressor of its own RNA transcript[2–4]. The properties of SV40 tsA mutants, which produce temperature-sensitive T antigen, provide evidence for another role of T antigen—in the initiation of viral DNA replication[5]. Mutations around the origin of DNA replication in SV40 lead to a *cis*-acting defect in DNA replication[6,7]. These mutations fall outside the region of the A gene which codes for T antigen but within the region known to bind T antigen[8] and the closely related and more easily prepared D2T protein[9,10] which is synthesized by an adeno–SV40 hybrid virus Ad2D2 and seems to differ from SV40 T antigen only in its amino-terminal segment[11]. Revertants of origin-defective mutants have been isolated with second site mutations within the DNA A gene coding sequence[12]. These data support the view that T antigen modulates transcription and replication of SV40 DNA by binding to one or more sites in and around the viral origin of replication. We present here evidence to show that mutants of SV40 carrying constructed mutations in the regulatory region[13] show reduced binding of D2T protein *in vitro*. This is the first demonstration in a eukaryote that a mutation in a regulatory DNA sequence alters the binding of a controlling protein *in vitro*.

Figure 1 shows the organization of the relevant part of the regulatory region in SV40. There are apparently three tandem D2T protein binding sites[9,10] spanning a segment known to be involved in the control of DNA replication[6,7,13,14]. The 5′ termini of the early mRNA and some of the late mRNAs also fall in this region of the DNA[15,16]. The sequence of this segment of wild-type DNA and of the three mutant DNAs used in this study are shown in Fig. 1. The mutants cs (cold-sensitive for plaque formation) 1088 and 1085 were constructed by site-specific mutagenesis[13]; their altered sequences fall within the region defined by DNase and methylation protection studies as the first D2T protein binding site[9,10]. By contrast, the mutant 812 lacks the second and third binding sites but the first site is unaltered[7]. These changes in sequence around the origin of replication allow distinction of the contribution of different parts of the regulatory region to T antigen binding.

An immunoassay for DNA-protein interaction has recently been developed[17]. An important feature of immunoassay procedures to study hormone–receptor interactions is that quantitative measurements of affinity can be made in the presence of complex protein mixtures. The ability to manipulate DNA suggests that DNA fragments of defined sequence can be used, like hormones, as probes for high-affinity DNA binding proteins in complex cell extracts. Assays of this sort are simple to perform, can be used to screen many samples such as mutant viruses and may be particularly valuable in studying the specific binding of regulatory proteins to chromatin rather than DNA. The defective adenovirus Ad2D2 synthesizes a protein related to SV40 T antigen[11]. This protein, D2T, is known to bind preferentially to DNA sequences around the SV40 origin of replication[10]. Using a monoclonal antibody (RMT-1) specific for SV40 T antigen, DNA fragments containing the regulatory sequences around the wild-type replication origin are preferentially precipitated. The controls in this experiment are of two types: first, extracts of HeLa cells infected with adenovirus Ad2$^+$ alone show no preferential precipitation of DNA fragments containing the SV40 origin. Second, SV40 DNA fragments lacking the origin of replication are only immunoprecipitated by partially purified D2T when the protein is present in excess and all the origin-containing fragments are already bound by D2T. The immunoassay was used to obtain both kinetic and equilibrium measurements of the affinity of DNA–protein interaction. Both the half life of the complex once formed and a Scatchard equilibrium treatment gave data consistent with an equilibrium dissociation constant of 10^{-12} M. Similar values have been obtained with the nitrocellulose filter binding assay[18].

Site-specific mutagenesis in the regulatory region of SV40 gives viruses with altered properties. The explanation proposed for this alteration in phenotype is that the interaction of T antigen with DNA is changed in the mutant virus. Consistent

0028-0836/81/090810—04$01.00

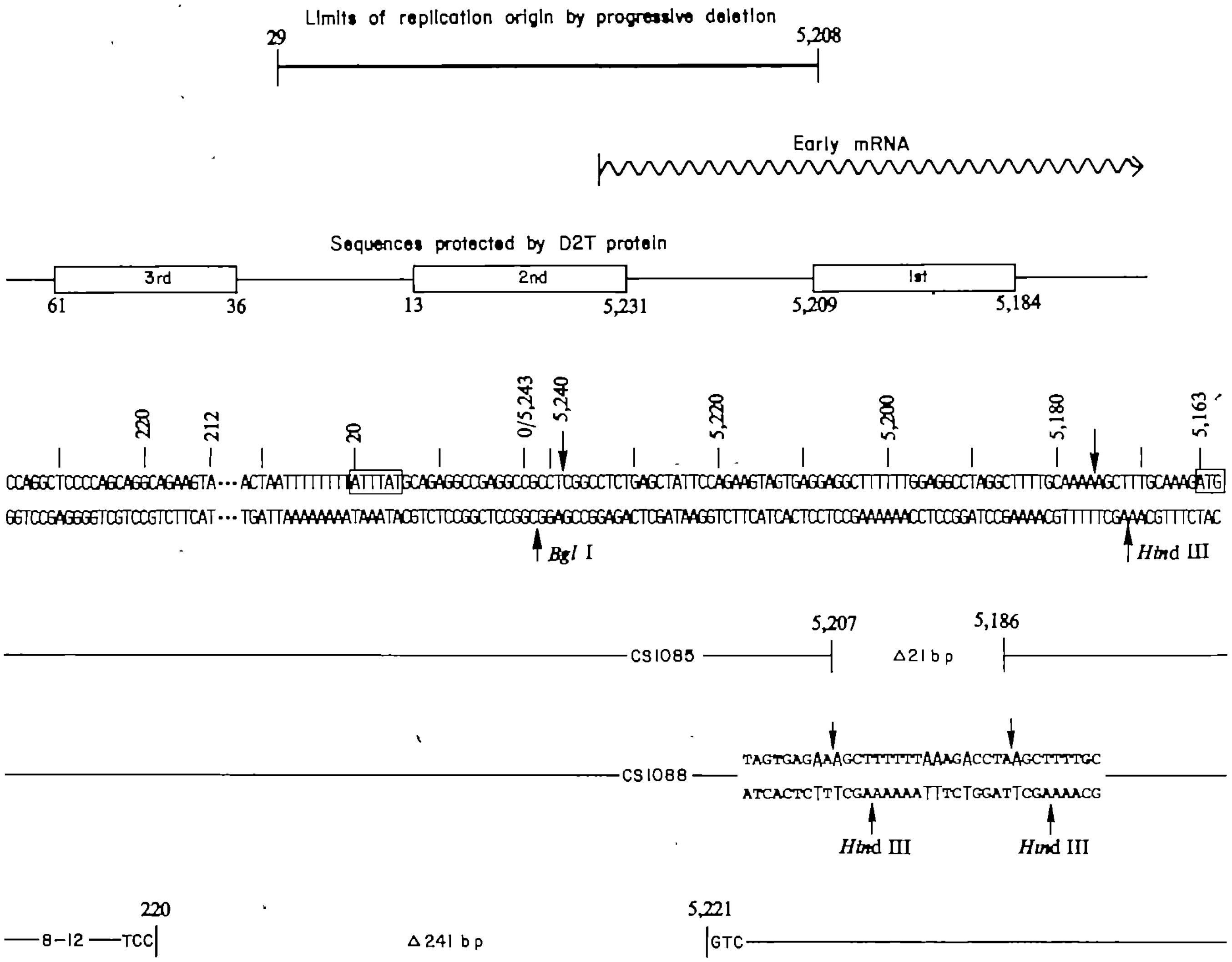

Limits of replication origin by progressive deletion
29
5,208
Early mRNA
Sequences protected by D2T protein
3rd
2nd
1st
61
36
13
5,231
5,209
5,184
220
212
20
0/5,243
5,240
5,220
5,200
5,180
5,163
CCAGGCTCCCCAGCAGGCAGAAGTA···ACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTTGCAAAGATG
GGTCCGAGGGGTCGTCCGTCTTCAT···TGATTAAAAAAAATAAATACGTCTCCGGCTCCGGCGGAGCCGGAGACTCGATAAGGTCTTCATCACTCCTCCGAAAAAACCTCCGGATCCGAAAACGTTTTTCGAAACGTTTCTAC
Bgl I
Hind III
5,207
5,186
CS1085
Δ21 bp
TAGTGAGAAAGCTTTTTTAAAGACCTAAGCTTTTGC
CS1088
ATCACTCTTTCGAAAAAATTTCTGGATTCGAAAACG
Hind III
Hind III
220
8-12
TCC
Δ241 bp
5,221
GTC

Fig. 1 The structure of the regulatory region of SV40. The wild-type DNA sequence around the origin of SV40 replication is shown in the middle of the figure (numbered according to the BB convention[1]). Arrows indicate cleavage sites for the restriction endonucleases *Hin*dIII and *Bgl*I. The initiation codon for large T and small t antigens and the ATTTAT box are enclosed on the early sense strand. Above this are shown the sequences protected by D2T from methylation and DNase I digestion[9,10], the 5′ end and direction of transcription of early mRNA[16] and the limits of the replication origin[13,14]. Below the wild-type sequence the structure of mutant DNAs used in this study are shown[7,13]. The base substitutions generated in cs 1088 by bisulphite treatment are shown in larger type. Note the two new *Hin*dIII sites in this mutant. b.p., Base pairs.

with this view, second-site revertants to wild-type phenotype have been mapped to the structural sequences coding for SV40 large T antigen[12]. Using a series of deletions around the origin the sequences responsible for the preferential immunoprecipitation of the regulatory region have been mapped within the limits of base pair 5 to base pair 5,189 (ref. 17). Here we have studied the specific D2T binding to virus genomes carrying less drastic alterations in their regulatory region. Our results show directly that these mutant viruses are altered in their ability to interact with a regulatory DNA binding protein.

When SV40 DNA is cut with the enzyme *Bst*NI (an isoschizomer of *Eco*RII, ref. 19), the origin of replication and the D2T protein binding sites are contained in the 311-base pair G fragment. After incubation of ^{32}P-labelled, *Bst*NI-digested wild-type SV40 DNA with D2T protein followed by immunoprecipitation with anti-D2T antibody, this fragment is preferentially recovered (Fig. 2). By contrast, *Bst*NI G fragments derived from cs 1088 and cs 1985 DNA and carrying mutations in the first binding site show a greatly reduced level of specific binding (Fig. 2). As the *Bst*NI G fragment of cs 1085 (containing a deletion of 21 base pairs) can be resolved from the wild-type G fragment by gel electrophoresis, mixtures of wild-type and mutant fragments could be tested for relative binding of D2T protein. The result confirms that D2T protein binds the mutant fragment markedly less than the wild-type fragment. As noted in Fig. 2 legend, the binding assays were carried out at a concentration of 20mM NaCl. At high concentrations of NaCl specific binding was inhibited, presumably due to inhibition of ionic interactions between D2T protein and the binding site, but the differential binding of mutant and wild-type G fragment was apparent at all salt concentrations tested.

The binding of the origin-containing DNA restriction fragment was measured relative to another restriction fragment. The autoradiograph shown in Fig. 2 was analysed using a computerized scanning system with four elements: (1) Optronix P1000 scanner; (2) PDP 11/60 computer; (3) Grinnel graphics system, and (4) the QUEST software system (J. Garrels, Cold Spring Harbor Laboratory). The ratio of the integrated intensity of the *Eco*RII/G fragments was 6.5 (s.d. ± 1.3) for wild-type DNA, 1.87 (s.d. ± 0.44) for the mutant cs 1085 and 0.81 (s.d. ± 0.25) for the mutant cs 1088. Note that the *Eco*RII/C fragment is 673 base pairs long and the Eco RII/G fragment only 311 base pairs, thus the relative molar binding of wild-type RII/G to RII/C fragments is 13.3. These figures clearly show that

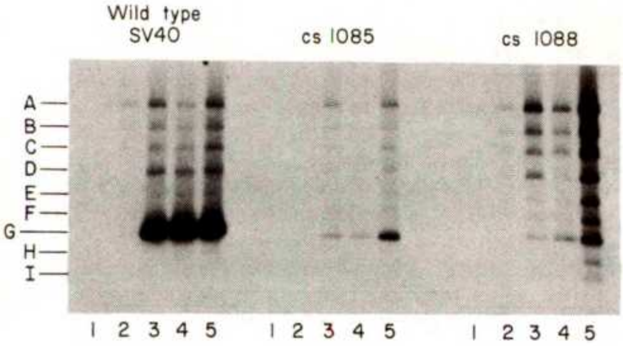


Fig. 2 Immunoprecipitation of SV40 DNA restriction fragments bound by D2T protein. 5 ng of *Bst*NI-digested SV40 or mutant DNA (specific activity 5×10^7 c.p.m. μg^{-1}) were incubated with increasing amounts of partially purified D2T protein: 1, 0 μl; 2, 25 μl; 3, 100 μl; 4, 200 μl; 5, 300 μl, in 20 mM sodium chloride, 2.0 mM dithiothreitol (DTT), 0.1% bovine serum albumin, 0.1 mm EDTA, 0.05% NP40, 40 $\mu g\ ml^{-1}$ phenylmethylsulphonyfluoride, 3% dimethyl sulphoxide, for 10 min at 25 °C. All the experiments reported in this paper were carried out with the same partially pure preparation of the D2T protein. The bound DNA fragments were then immunoprecipitated by addition of 30 μl of monoclonal anti-D2T antibody for 10 min, followed by 20 μl of rabbit anti-mouse immunoglobulin and 60 μl of fixed *Staphyloccocus aureus*. The detailed characterization of this assay is described elsewhere[17]. The immunoprecipitated fragments were released from the fixed bacteria with 1% SDS in 10 mM Tris-HCl *p*H 8.0/1 mM EDTA and analysed by agarose gel electrophoresis and autoradiography. Fragment G contains the replication origin and adjacent sequences. The D2T protein used in this assay was partially purified from HeLa cells infected with $Ad2^+D2$ virus (50 plaque-forming units per cell) and grown at 37 °C for 30 h. Nuclei were prepared by hypotonic lysis and the nuclear D2T protein extracted in buffer A (0.15 M NaCl, 10 mm Tris-HCl *p*H 8.0, 1 mm EDTA, 0.5 mm DTT). This crude nuclear fraction was further purified by chromatography on DE52 (elution in 0.25 m NaCl, 10 mm Tris-HCl *p*H 8.0, 1 mm EDTA, 0.5 mm DTT). This fraction was then precipitated with 50% ammonium sulphate, dissolved in buffer A plus 30% glycerol and stored at −70 °C. Polyacrylamide gel electrophoresis showed D2T to be one of four major proteins present at this stage.

quantitative scanning of the autoradiograph in Fig. 2 confirms that mutations in the first D2T binding site substantially reduce the D2T-mediated precipitation of these DNA fragments.

In Figure 3 the binding of the *Hin*dIII fragments of wild-type SV40, cs 1088 and 812 are compared. The base substitution in cs 1088 generate two new *Hin*dIII restriction sites close to the start of the second D2T protein binding segment[13]. These new restriction sites allow the preparation of a DNA fragment containing only binding segments 2 and 3; as seen in Fig. 3 this fragment shows only very slight preferential binding to D2T protein. In contrast, the wild-type and 812 *Hin*dIII fragments bind D2T protein strongly. Note that in the case of mutant 812 this strong binding occurs in the absence of the second and third sites. We can conclude that the first site alone can bind D2T protein tightly, but the second and third binding sites are not strong independent sites.

Our main conclusion from these studies is that mutations in a regulatory segment of the SV40 genome, outside structural gene sequences, affect the interaction of this segment of viral DNA with a regulatory protein. Mutations located in the previously defined first D2T protein binding site within the regulatory segment[9,10] cause a marked decrease in binding of the D2T protein. (Recently we have found that these mutations also cause diminished binding of T antigen partially purified from SV40-infected monkey cells.) The decreased binding is not related to the use of monoclonal antibody in our assay, as serum from hamsters bearing SV40-induced tumours served equally well. In contrast, mutations in the second T antigen binding site had little or no effect on binding of D2T protein (Fig. 3 and ref. 17). Thus it seems that D2T protein can bind independently to the first site, but poorly or not at all to the isolated second plus third sites. Testing of other mutants with base pair substitutions in this region should help determine which bases contribute most to the specific interaction with D2T protein and with T antigen.

The cold-sensitive regulatory mutants used in this study form small plaques at 32 °C but wild-type plaques at 37 and 40 °C (ref. 13). They therefore seem to have a defect in virus growth at the lower temperature, previously postulated to be due to cold-sensitive interaction of a regulatory protein with an altered signal in the viral DNA[13]. We attempted to demonstrate a differential effect of temperature on the binding of D2T protein to mutant DNA fragments, but due to greater non-specific binding at elevated temperatures, a clear effect was not found.

Regarding the possible physiological significance of the interaction of T antigen in the first binding site of the SV40 regulatory segment, an attractive hypothesis is that this interaction is responsible for inhibiting transcription of the gene for T

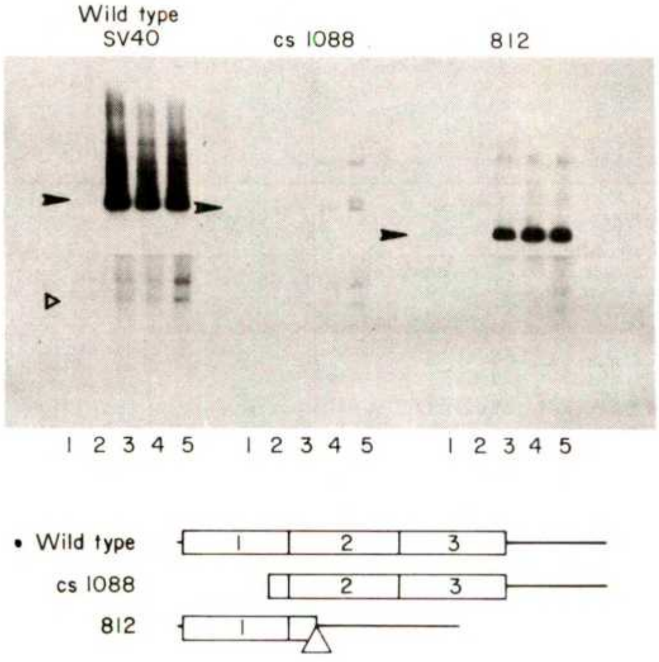


Fig. 3 Immunoprecipitation of *Hin*dIII DNA fragments containing different combinations of the three D2T protein binding sites. The structure of the *Hin*dIII fragments spanning the origin is shown in the lower half of the figure with D2T binding sites represented by numbered boxes. The wild-type *Hin*dIII fragment starts close to the first site and spans all three sites. The *Hin*dIII fragment from deletion mutant 812 contains only an intact first site. The *Hin*dIII fragment derived from cs 1088 contains only sites 2 and 3. The D2T binding properties of these fragments were assayed as described in Fig. 2, using 100 μl of D2T per point. The solid arrowheads indicate the position of the origin-containing fragments after immunoprecipitation and electrophoresis; the open arrowhead shows a smaller *Hin*dIII restriction fragment of SV40 used as an internal control in computerized scanning of the autoradiograph. The lower half of the figure has been overexposed to show the control fragments. Integrated intensity in the origin fragment relative to the control fragment was 10.75 for wild-type DNA, 4.39 for the mutant cs 1085 and 8.79 for the deletion mutant 812 which lacks the second and third binding sites. The high binding of fragments larger than the origin-containing fragment of wild-type SV40 DNA is caused by a small proportion of partial restriction fragments. These immunoprecipitated partial fragments were not included in the relative binding figures given.

antigen. Consistent with this possibility are the presently known properties of mutants cs 1085 and cs 1088, that is, overproduction of early viral RNA and T antigen, and nearly normal rates of viral DNA replication (D. D. and D. Nathans, unpublished observations). More detailed studies of these and similar mutants, including *in vitro* transcription of their DNAs, will be required to test this hypothesis fully.

We thank Yakov Gluzman for DNA from mutant 812, Joe Sambrook and Daniel Nathans for support and critical reading of this manuscript. This research was supported by grants from the US NCI (2 P01 CA 16519 and CA 13106-09). D.D. is a trainee under the Medical Scientists Training Program of the US National Institutes of General Medical Sciences (grant no. GM 07309). During part of this work R.M. was a Research Fellow of the MRC.

Received 4 August; accepted 17 December 1980.

1. Tooze, J. *Molecular Biology of Tumor Viruses* 2nd edn, part 2 *DNA Tumor Viruses* (Cold Spring Harbor laboratory, New York, 1980).
2. Tegtmeyer, P., Schwarz, M., Collins, J. K. & Rundell, K. *J. Virol.* **16,** 168–178 (1975).
3. Reed, S. L., Stark, G. S. & Alwine, S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3083–3087 (1976).
4. Khoury, G. & May, E. *J. Virol.* **23,** 167–176 (1978).
5. Tegtmeyer, P. *J. Virol.* **15,** 613–618 (1975).
6. Shortle, D. & Nathans, D. *J. molec. Biol.* **131,** 801–817 (1979).
7. Gluzman, Y., Frisque, R. J. & Sambrook, J. F. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3898–3902 (1980).
8. Shalloway, D., Kleinberger, T. & Livingston, D. M. *Cell* **20,** 411–422 (1980).
9. Tjian, R. *Cell* **13,** 165–179 (1978).
10. Tjian, R. *Cold Spring Harb. Symp. quant. Biol.* **43,** 655–662 (1978).
11. Hassell, J. A., Lukanidin, E., Fey, G. & Sambrook, J. *J. molec. Biol.* **120,** 209–247 (1978).
12. Shortle, D., Margolskee, R. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6128–6131 (1979).
13. DiMaio, D. & Nathans, D. *J. molec. Biol.* **140,** 129–142 (1980).
14. Subramanian, K. & Shenk, T. *Nucleic Acids Res.* **5,** 3635–3642 (1978).
15. Ghosh, P. K., Reddy, V. B., Swinscoe, J., Lebowitz, P. & Weissman, S. M. *J. molec. Biol.* **126,** 813–846 (1978).
16. Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. *J. Virol.* **30,** 279–296 (1979).
17. McKay, R. *J. molec. Biol.* **145,** (in the press).
18. Jessel, D., Landau, T., Hudson, J., Tenen, D. & Livingston, D. M. *Cell* **8,** 525–545 (1976).
19. Roberts, R. J. *Nucleic Acids Res.* **8,** 463–680 (1980).

Network structure in gels of rod-like polypeptides

Kohji Tohyama & Wilmer G. Miller

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA

The formation of a mechanically self-supporting, macromolecular gel or network is understood on a molecular scale in terms of cross-links or branch points, either of a permanent nature by covalent bond formation, or of a reversible nature, such as in the gelation of gelatin. In thermally reversible gels each branch point may contain numerous monomeric units, indicative of local crystallite or aggregate formation, as in cellulose acetate[1] or gelatin[2–4], or only two monomeric units as in polyacrylylglycinamide[5]. The formation of reversible networks in nonionic, rod-like polypeptide homopolymers[6–8], particularly at low concentrations (<0.1 wt%), is surprising in that the molecular origin of the branch points is not obvious. We have suggested previously[8] that this network constitutes a thermodynamic phase resulting from a particular kinetic mechanism of phase formation made favourable by the unusual polymer–diluent phase equilibria occurring with stiff-chain polymers. Here, the network is visualized by electron microscopy, and shown to be compatible with the proposed mechanism.

Examples of gels visualized by microscopy are shown in Fig. 1. Typically, an isotropic solution was cooled to the desired gelation temperature, held overnight, quenched in Freon 22 cooled with liquid nitrogen, and at −150 °C fractured, etched, shadowed (Pt–C) and replicated with a thick carbon layer. The replica, after removal of polymeric material, was viewed in a Zeiss 9 S-2 electron microscope. Polybenzylglutamate (MW 200,000) networks generally appeared random (Fig. 1*a*); occasionally regions seemed to have been under uniaxial deformation when quenched (Fig. 1*b*). Virtually all strands begin and end at branch points.

A freshly fractured surface was shadowed without etching. A typical micrograph (Fig. 1*c*) shows many ends and spurs, as one might expect with a fractured network. Occasionally a fracture-knife gouge across a network strand was clearly visible. Furthermore, micrographs from isotropic solutions cooled rapidly through the gelation region by quenching in liquid nitrogen showed no network structure. These results indicate that the visualization of the polypeptide networks was not an artefact of sample preparation.

For comparison a network formed from a more flexible polymer was visualized—a gelatin network (Fig. 1*d*). The micrograph is quite distinct from the polypeptide micrographs, but similar to micrographs of agar and acrylamide gels prepared in a similar manner[9,10]. Micrographs of other gels have been obtained using different preparation techniques[2,11,12]. However, we show elsewhere[13] that our gelatin micrograph is an artefact of sample preparation.

The micrographs of the polypeptide networks reveal strands or fibres with cross-links that do not result from the contact of nonaligned fibres, that is, the cross-links/branch points tend not to have four arms. The fibres are of diameters ranging from hundreds to more than a thousand angstroms, and are composed of bundles of aligned rods (rod diameter ~20Å). The spacing between fibres seems to be typically of the order of 1 μm. These values are not certain as the platinum may shadow polymer and solvent differently in the unetched samples, and a depth scale is lacking in the etched samples. However, etched and unetched samples do give similar results.

We believe that the origin and structure of these networks may be the result of the kinetics of phase separation. Typically, when a solution is brought from a uniphasic into a biphasic region the new phase forms by a nucleation and growth mechanism. However, if the solution is brought into a region of thermodynamic instability ($\sigma^2G/\sigma c^2 < 0$ where G = Gibbs free energy and c = concentration) rather than metastability, the spinodal decomposition theory of Cahn and Hilliard[14] predicts phase separation by a non-nucleation mechanism, whereby the spatial variation in solution composition is continuous with a characteristic periodicity. Material flows spontaneously against a concentration gradient. Early in the phase separation each phase is predicted to be continuous in the other phase. In systems which flow, ageing occurs by tube flow[15], leading eventually to the breakup of the network-like morphology. Most aspects of the theory have been tested experimentally[15–21].

Application to the polypeptide gels is made by reference to the Flory phase diagram for rigid rod polymers (Fig. 2) which corresponds closely to that observed for polybenzylglutamate[7,22]. A network forms when a solution is cooled and crosses into the wide biphasic region (designated 'Network'). The very drastic polymer–solvent unmixing which occurs with only a slight change in polymer–solvent interaction should allow entry into the unstable region even at low concentrations. Thus, if phase separation occurs by spinodal decomposition, the network which forms is simply the polymer-rich phase, a damp fibre, which is bicontinuous with the solvent-rich phase. The polymer-rich phase has no tendency to flow, which accounts for the persistence of the network after phase separation is complete. The mean periodicity in concentration fluctuation can be predicted[18,19,23], which should correspond to the interfibre spacing. Fibre thickness may then be calculated by mass balance. Application to our samples predicts a spacing of the order of 1 μm and a mean fibre thickness of several hundred angstroms[23], which agree with our observed values.

Other rod-like synthetic polymers form gels[20]. Also, many biological macromolecules form rod-like assemblies and gels, which sometimes affect cell morphology, for example, sickle-cell

0028–0836/81/090813—02$01.00

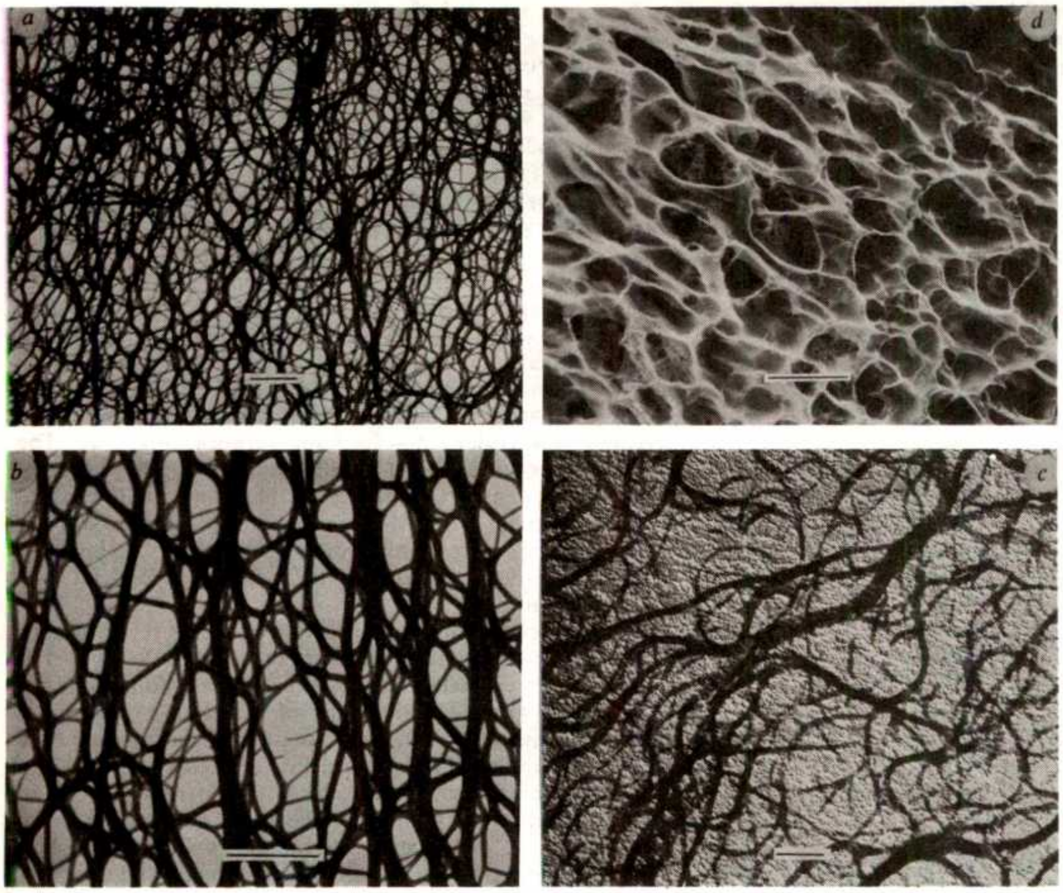


Fig. 1 Electron micrographs of 1 wt% polybenzylglutamate gels (*a–c*) and of a 3 wt% gelatin gel (*d*). Scale bar 1 μm. Etch times: 10 s (*a*, *b*), 0 s (*c*), 60 s (*d*).

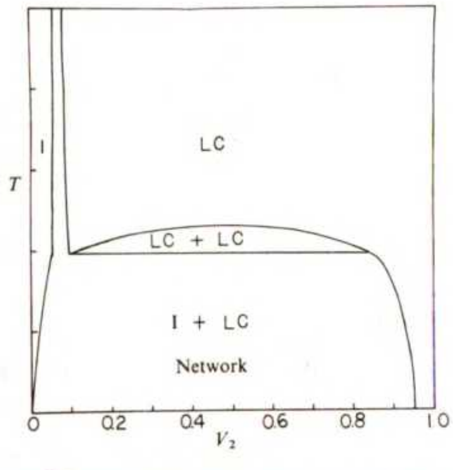


Fig. 2 Flory[22] temperature–composition (volume fraction polymer) phase diagram for rigid rod–diluent (rod axial ratio 150), which corresponds closely to that observed for polybenzylglutamate[7]. A network (gel) forms when a solution is cooled and crosses into the wide biphasic region (Network). The phases present are isotropic (I) and liquid crystalline (ordered) (LC).

haemoglobin, actin, tubulin and tobacco mosaic virus[20,24–26]. Thus the mechanism by which a rod-like polymer forms a network spanning the confining vessel may be of interest in understanding passive structure formation in more complex systems.

We thank the NIH for funding.

Received 19 September 1980; accepted 16 January 1981.

1. Goebel, K. D. & Berry, G. C. *J. Polym. Sci.* pt A-2 **15**, 555–577 (1977).
2. Boedtker, H. & Doty, P. M. *J. phys. Chem.* **58**, 968–983 (1954).
3. Veis, A. *The Macromolecular Chemistry of Gelatin* (Academic, New York, 1964).
4. Rogovins, L. Z. & Slonimskii, G. L. *Usp. Khim.* **43**, 1102–1135 (1974).
5. Haas, H. C., Chiklis, C. K. & Moreau, R. D. *J. Polym. Sci.* A-1 **8**, 1131–1145 (1970).
6. Doty, P. M., Bradbury, J. H. & Holtzer, A. M. *J. Am. chem. Soc.* **78**, 947–954 (1956).
7. Miller, W. G. *et al. Pure appl. Chem.* **38**, 37–58 (1974).
8. Miller, W. G., Kou, L. J., Tohyama, K. & Voltaggio, V. *J. Polym. Sci. Polym. Symp.* **65**, 91–106 (1978).
9. Steere, R. L. & Erbe, E. F. *Proc. 35th A. Electron Microscope Soc. Am.* (ed. Bailey, G. W.) 606–608 (Claitors, Baton Rouge, 1977).
10. van Emmerik, P. T., Smolders, C. A. & Geymayer, W. *Eur. Polym. J.* **9**, 309–314 (1973).
11. Mühlethaler, K. *Makromolek. Chem.* **2**, 143–171 (1948).
12. Theimer, W. *Z. Naturforsch* **15B**, 346–350 (1960).
13. Tohyama, K. & Miller, W. G. (in preparation).
14. Cahn, J.W. & Hilliard, J. E. *J. chem. Phys.* **31**, 688–699 (1959); **42**, 93–99 (1965).
15. McMaster, D. L. P. in *Copolymers, Polyblends and Composites* 43–65 (Am. chem. Soc., Washington DC, 1975).
16. Goldsbrough, J. *Sci. Prog.* **60**, 281–297 (1972).
17. Jantzen, C. M. F. & Herman, H. in *Phase Diagrams* Vol. 5 (ed. Alper, A. M.) Ch. 3 (Academic, New York, 1978).
18. van Aartsen, J. J. *Eur. Polym. J.* **6**, 919–924 (1970).
19. Smolders, C. A., van Aartsen, J. J. & Steenberger, A. *Kolloid Zh.* **243**, 14–20 (1971).
20. Miller, W. G. *A. Rev. phys. Chem.* **29**, 519–535 (1978); see refs therein.
21. Pines, E. & Prins, W. *Macromolecules* **6**, 888–895 (1973).
22. Flory, P. J. *Proc. R. Soc.* A**234**, 73–89 (1956).
23. Chakrabarti, S., Tohyama, K. & Miller, W. G. (in preparation).
24. Oosawa, F. & Asakura, S. *Thermodynamics of the Polymerization of Protein* (Academic, New York, 1975).
25. Bawden, F. C. & Pirie, N. W. *Proc. R. Soc.* B**123**, 274–320 (1937).
26. Hartwig, J. H. & Stossel, T. P. *J. molec. Biol.* **134**, 539–553 (1979).

Structure of the protein and DNA in fd filamentous bacterial virus

D. W. Banner, C. Nave & D. A. Marvin

European Molecular Biology Laboratory, Postfach 10.2209,
6900 Heidelberg, FRG

The virion of filamentous bacterial viruses comprises a cylindrical protein shell of o.d. ~60 Å and i.d. 20 Å, containing a single-stranded circular DNA molecule which has two oppositely directed but not base-paired strands extending the length of the virion. The assembly of the virion involves an intracellular prepackaging of the DNA with a viral DNA-binding protein which is then displaced by the coat protein as the

0028-0836/81/090814—03$01.00

growing virion crosses the bacterial membrane[1]. Studies of the virion by X-ray fibre diffraction show that the protein coat consists largely of α-helices oriented roughly parallel to the axis of the virion[2,3]. As the normal to a planar peptide tends to align normal to a magnetic field, it is possible to improve significantly the orientation of virions in fibres using a strong magnet[4,5]. The success of this technique with the Pf1 strain of virus led us to apply it to the better-known fd (f1, M13) strain. We report here new information about the arrangement of protein and DNA in the fd virion obtained from the improved diffraction pattern (Fig. 1).

The fd virion is ~8,700 Å long; its DNA contains 6,408 nucleotides, and its coat contains ~2,760 50-residue protein molecules. The distribution of intensity on the fd fibre patterns indicates that rods of electron density slope from large to small radius and also slew radially, as proposed[6] for the similar virus Pf1. The position of layer lines on the diffraction patterns gives information about the relative positions of the protein subunits. There is a set of layer lines at orders of 32 Å, and a set of meridional reflections at orders of 16 Å. Calculation of mass per length (as described in Table 9 of ref. 7) using accurate molecular weight data derived from the DNA sequence[8], shows that in the 16-Å repeat distance there are 5.1 protein subunits (±7%) and 11.8 nucleotides (±4%). These numbers imply 2.72 Å average axial separation between nucleotides in each of the two DNA strands.

The meridional reflection at 16 Å was previously interpreted[3] as being due to a periodic perturbation of a helical arrangement of subunits, with a repeat of the perturbation every five subunits. This perturbation predicts satellite layer lines which could not be seen on poorly oriented patterns, but their possible presence could not be ruled out. These satellite layer lines are clearly absent on Fig. 1, therefore such a perturbation is not supported by the data. An alternative interpretation[9,10] of the 16-Å meridional reflection is that the virion symmetry involves a 5-fold axial rotation axis combined with a 2-fold axial screw axis of pitch 32 Å (Fig. 2). The difference between this kind of arrangement and a perturbed helix is small (see ref. 11, Fig. 10). In particular, subunits with the same shape as in the perturbed helix model pack well (Fig. 3) in the symmetry of Fig. 2, and give a calculated diffraction pattern which agrees with the strong 10-Å equatorial and near-equatorial observed data (Fig. 4). The local interactions between nearest neighbours shown in Fig. 3 are similar to those of the structurally related virus Pf1 (ref. 10; Fig. 3) at 7 Å resolution.

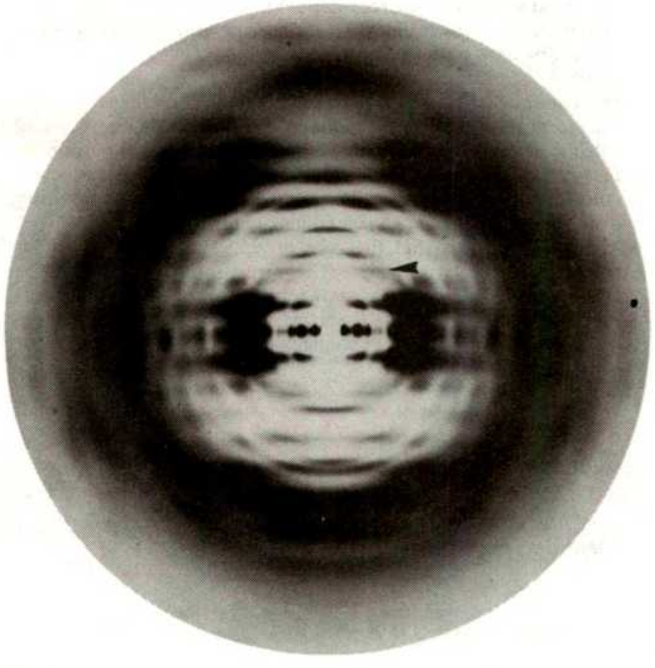


Fig. 1 X-ray fibre diffraction pattern of fd filamentous virus. Fibre axis vertical. Preparation of virus and X-ray fibre diffraction techniques are described elsewhere[3]. In the purified virus solution, >98% of the protein migrated with the major coat protein and >95% of the DNA migrated with viral DNA in appropriate gel electrophoreses. The virus solution was titrated to *p*H 2 (below the isoelectric point) before preparing fibres, because fibres prepared from virus solutions at neutral *p*H give double orientation (A. Fowler, personal communication). Protein layer lines occur at orders of $c = 32.3$ Å with meridional reflections on even layer lines. DNA layer lines occur at 26.7, 13.35 (arrow), 3.42 and 3.02 Å. Unit cell parameter $a = 60.7$ Å; 92% relative humidity; specimen–film distance 4.20 cm; tilt 16.6°; standard deviation of disorientation 3.5°. Fibre fd 45.2; film 5148. ×1.6.

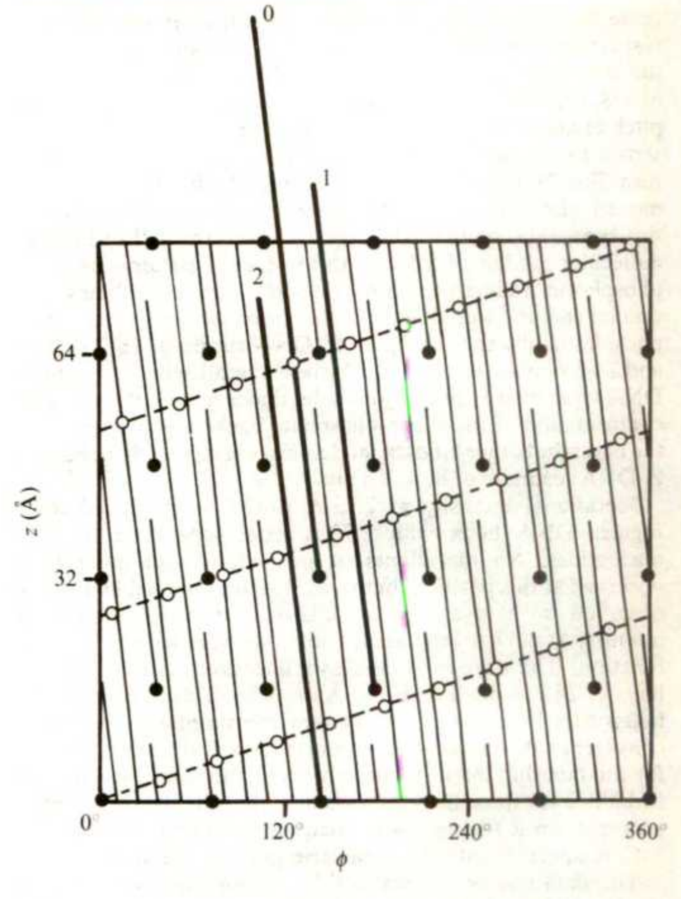


Fig. 2 Surface lattice of the fd virion. The surface lattice is generated by projecting equivalent points along radii onto a cylinder of 13 Å radius, and then cutting the cylinder along a vertical line to give a 2-dimensional representation. ●, Projections of equivalent points in each protein subunit, and the solid lines represent the projections of the α-helix axes. The bold solid lines indicate the three nearest neighbours shown in Fig. 3. ○, Projections of equivalent points in each DNA nucleotide pair, and the broken line represents the projection of one strand of the DNA helix. The fact that the DNA helix passes through surface lattice points of the protein at a (ϕ, z) of (0, 0), (72°, 32 Å), and (144°, 64 Å) may be fortuitious, but note that a nucleotide/protein ratio of 2.30 would give identical DNA–protein environments every 23 nucleotides along one DNA strand (every 64 Å along the axis), and a nucleotide/protein ratio of 2.40 would give identical DNA–protein environments every 12 nucleotides (every 32 Å). The relative phase of the DNA and protein is unknown; that is, the DNA helix could be shifted in ϕ and/or z relative to the protein helix. The line-group symmetry[14] of the protein assembly is **sr**, $N = 5$.

The diffraction patterns of magnetically aligned fd often show faint layer lines at the first two orders of a 26.7-Å periodicity. If these layer lines are attributed to J_1 and J_2 Bessel functions respectively, the radial positions of the intensity maxima along the layer lines would correspond to electron density at a radius of ~8.5 Å. We interpret these layer lines as being due to the pitch of helical viral DNA. DNA of this pitch with 2.72 Å axial separation between nucleotides would then have 9.8 units per turn (Fig. 2). If there is no base pairing, the two strands can be moved relative to one another. We know of no DNA structure of this type, thus we used the approach (see ref. 12) of building molecular models of DNA with these helix parameters and a phosphorus radius of 8.5 Å. (Models with phosphates at a smaller radius[13] would imply an abnormally short phosphorus–phosphorus distance along an fd DNA strand of 26.7 Å pitch and 2.72 Å axial separation.) The best overall fit to the first three DNA layer lines is given by a model that is unexceptional in the conformation of the sugar-phosphate backbone and the tilt of the bases, but other kinds of models (for example, left-handed or Z-DNA) cannot yet be ruled out.

Meridional intensity at 2.72 Å would be predicted for a regular DNA helix with 2.72 Å axial separation between nucleotides. No meridional or near-meridional intensity is observed at this position, but weak near-meridional intensity is observed on a layer line at 3.02 Å, and strong meridional intensity at 3.42 Å (apparently not consistent with a J_2 Bessel function). These are two of the layer lines predicted for a regular helix of 26.7 Å pitch and 2.72 Å unit rise. These observations, taken together with UV dichroism measurements, would be consistent with a tilted-base conformation of the DNA[7], except for the fact that intensity directly on the meridian at 3.42 Å is forbidden for these helix parameters. Thus the bases are not all arranged on a simple helix—many are stacked axially about 3.42 Å apart despite the mean rise per nucleotide of 2.72 Å.

The data can be summarized by identifying three separate clusters of layer lines on the diffraction pattern. The layer lines with a 32-Å repeat are due to coat protein molecules; the layer lines with a 26.7-Å repeat are due to the DNA which, at low resolution, follows a helix of 26.7 Å pitch and 8.5 Å radius; and the 3.42 Å meridional reflection is indicative of base stacking but is not consistent with a regular helix of 2.72 Å rise per nucleotide.

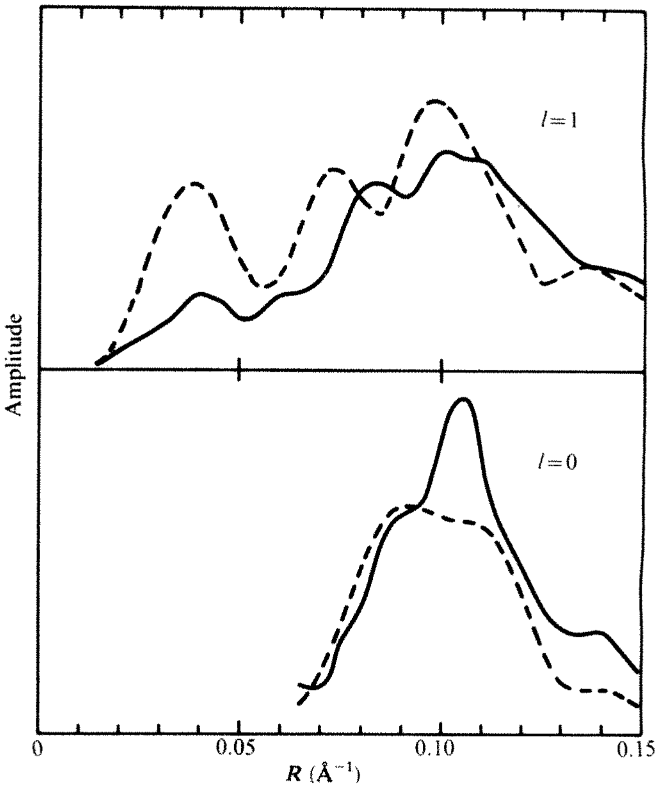


Fig. 4 Distribution of amplitude (absolute values) along layer lines in the strong 10-Å equatorial and near-equatorial region. Observed amplitudes are represented by solid lines, and calculated amplitudes by broken lines. Intensity data were measured[6] from the diffraction pattern shown in Fig. 1. The sharp peak at 0.105 Å^{-1} for $l = 0$ is due to crystalline reflections superposed on the continuous transform. The cylindrically averaged Fourier transform was calculated[6] for the model of Fig. 3. Observed and calculated intensities due to protein were scaled together, using the first seven layer lines, to $R = 0.15$ Å^{-1}. The J_0 Bessel function contribution is not shown. The similarity between the intensity distribution on the fd 32-Å layer lines and the layer lines of Pf1 (ref. 4; Fig. 3) suggests that the coat protein subunit has a similar structure in the two virions. This structure could be either a continuous α-helix rod[10] as shown in Fig. 3, or two α-helix segments occupying roughly the same space in the virion as the continuous rod, but connected in a different way.

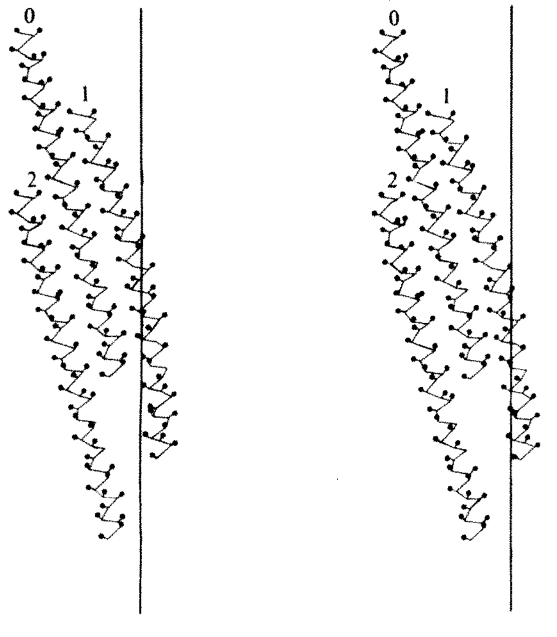


Fig. 3 Stereo pairs showing the interactions between an α-helical subunit and its nearest neighbours in a model of the fd virion fitting the data to 7 Å resolution. The lines connect α-carbons along the sequence. The points represent positions of β-carbon atoms. The view is perpendicular to the virion axis (vertical line) with the N-terminus upwards. These three molecules correspond to the bold lines in Fig. 2.

Received 2 September; accepted 15 December 1980.

1. Denhardt, D. T., Dressler, D. & Ray, D. S. *The Single-Stranded DNA Phages* (Cold Spring Harbor Laboratory, New York, 1978).
2. Marvin, D. A. *J. molec. Biol.* **15,** 8–17 (1966).
3. Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J. & Marvin, F. J. *J. molec. Biol.* **88,** 581–600 (1974).
4. Nave, C. *et al. Nature* **281,** 232–234 (1979).
5. Torbet, J. & Maret, G. *J. molec. Biol.* **134,** 843–845 (1979).
6. Marvin, D. A. & Wachtel, E. J. *Nature* **253,** 19–23 (1975); *Phil. Trans. R. Soc.* **B276,** 81–98 (1976).
7. Marvin, D. A. & Hohn, B. *Bact. Rev.* **33,** 172–209 (1969).
8. Beck, E. *et al. Nucleic Acids Res.* **5,** 4495–4502 (1978).
9. Makowski, L. & Caspar, D. L. D. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 627–643 (Cold Spring Harbor Laboratory, New York, 1978).
10. Marvin, D. A. *et al.* in *Structural Aspects of Recognition and Assembly in Biological Macromolecules* (Elsevier, Amsterdam, 1981).
11. Folkhard, W. *et al. J. molec. Biol.* **130,** 145–160 (1979).
12. Zimmerman, S. B. *Biopolymers* **16,** 749–763 (1977).
13. Day, L. A., Wiseman, R. L. & Marzec, C. J. *Nucleic Acids Res.* **7,** 1393–1403 (1979).
14. Klug, A., Crick, F. H. C. & Wyckoff, H. W. *Acta crystallogr.* **11,** 199–213 (1958).

Structure of TANDEM and its implication for bifunctional intercalation into DNA

M. A. Viswamitra*, Olga Kennard†, William B. T. Cruse†, Ernst Egert†, George M. Sheldrick‡, Peter G. Jones‡, Michael J. Waring§, Larry P. G. Wakelin§ & Richard K. Olsen‖

* Indian Institute of Science, Bangalore, India
† University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK
‡ Anorganisch-Chemisches Institut, University of Göttingen, FRG
§ Department of Pharmacology, Medical School, University of Cambridge, Cambridge CB2 2QQ, UK
‖ Department of Chemistry and Biochemistry, Utah State University, Logan, Utah 84321, USA

Quinoxaline antibiotics (Fig. 1*a, b*) form a useful group of compounds for the study of drug–nucleic acid interactions[1,2]. They consist of a cross-bridged cyclic octadepsipeptide, variously modified, bearing two quinoxaline chromophores. These antibiotics intercalate bifunctionally into DNA[2,3] probably via the narrow groove, forming a complex in which, most probably, two base pairs are sandwiched between the chromophores[4,5]. Depending on the nature of their sulphur-containing cross-bridge and modifications to their amino acid side chains, they display characteristic patterns of nucleotide sequence selectivity when binding to DNAs of different base composition and to synthetic polydeoxynucleotides[4,6,7]. This specificity has been tentatively ascribed to specific hydrogen-bonding interactions between functional groups in the DNA and complementary moieties on the peptide ring[2,4,5]. Variations in selectivity have been attributed both to changes in the conformation of the peptide backbone[6] and to modifications of the cross-bridge[7]. These suggestions were made, however, in the absence of firm knowledge about the three-dimensional structure and conformation of the antibiotic molecules. We now report the X-ray structure analysis of the synthetic analogue of the antibiotic triostin A, TANDEM (des-*N*-tetramethyl triostin A) (Fig. 1*c*), which binds preferentially to alternating adenine–thymine sequences[7]. The X-ray structure provides a starting point for exploring the origin of this specificity and suggests possible models for the binding of other members of the quinoxaline series.

Hydrated crystals of TANDEM, $C_{46}H_{54}N_{12}O_{12}S_2.12H_2O$, were obtained from a mixture of dimethyl formamide/ethanol/water. Crystal data: orthorhombic, $P2_12_12_1$, $Z=4$, $a=18.400$, $b=22.126$, $c=16.435$ Å. Diffractometer measurements for 5,422 independent reflections were obtained with MoKα radiation to a resolution of 0.85 Å at 7 ± 1 °C. The structure was solved with some difficulty by a combination of direct methods and a novel iterative Fourier procedure[8]. Refinement, with all non-hydrogen atoms anisotropic, converged at $R=14.7\%$ for 4,111 data with $F>2\sigma(F)$. Full details will be published elsewhere.

Figure 2 shows the molecular structure of TANDEM. The octadepsipeptide ring is held in a relatively rigid conformation through intramolecular hydrogen bonds NH(L-Val)... O(L-Ala). The quinoxaline chromophores and the disulphide cross-bridge project on opposite sides of the peptide ring. The carbonyl groups of L-Cys and L-Val also project from the side opposite to the chromophores. The molecule displays a helical twist and has an approximate 2-fold axis of symmetry. Bond lengths and angles fall within the normal ranges.

The mean planes of the two chromophores are inclined at an angle of 14°. The distance between the Ser α-carbon atoms is 12.2 Å. Assuming that the intercalated complex is of the Fuller-Waring type[9], with minimal distortion from a classical B-DNA helix, we would expect the chromophore planes to lie parallel, approximately orthogonal to a line joining the Ser α-carbon atoms and separated by a perpendicular distance of 10–11 Å, to allow for bifunctional intercalation[10]. These geometrical constraints are readily imposed on the molecule derived from the X-ray analysis by rotations of approximately 20° about the C_α–N bonds of D-Ser. Such rotations do not result in steric hindrance with neighbouring groups[5,11]. The resultant structure (Fig. 3) can be fitted into the narrow groove of a B-DNA helix,

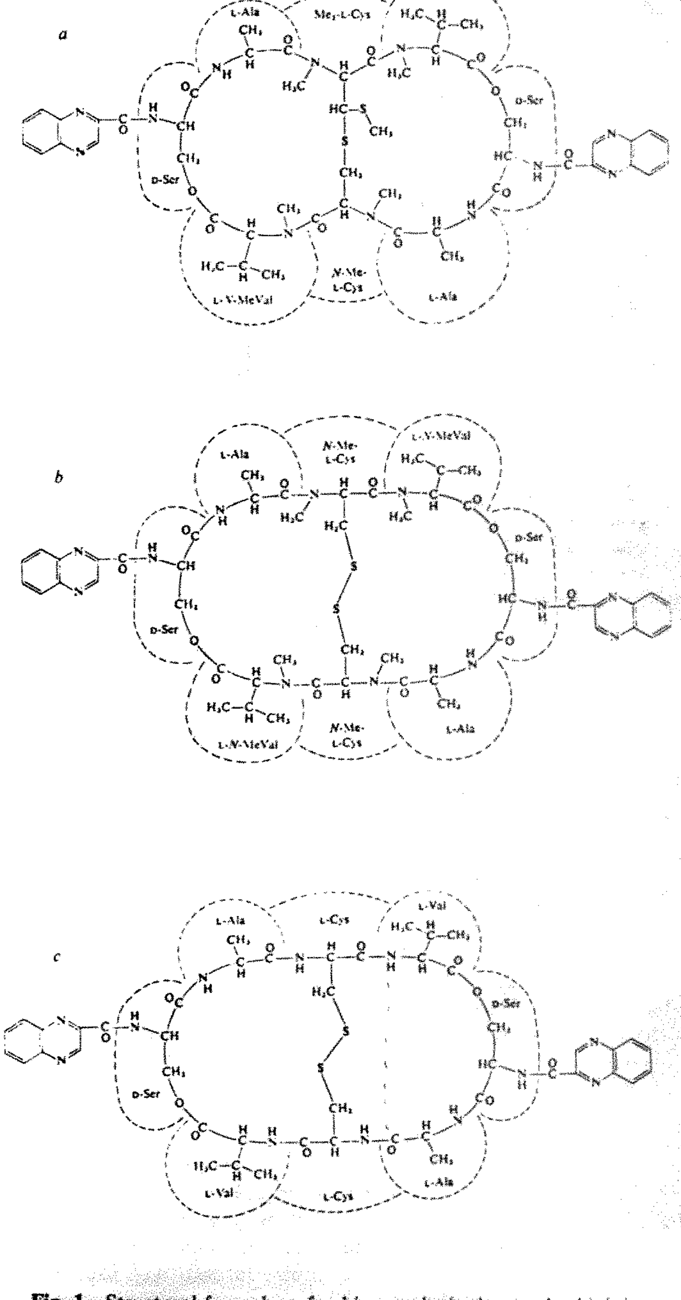


Fig. 1 Structural formulae of echinomycin (quinomycin A) (*a*), triostin A (*b*), TANDEM (des-*N*-tetramethyl triostin A) (*c*).

0028-0836/81/090817—03$01.00

which is known from solution experiments to be unwound by 45.5° on intercalation of TANDEM[2,7] (based on an unwinding angle of 26° for ethidium[12]). In this bis-intercalated model the disulphide cross-bridge projects outwards away from the helix, two base pairs are sandwiched between the chromophores and various functional groups of the cyclic peptide are placed in the vicinity of potential hydrogen-bonding sites on the DNA bases. The pseudo-2-fold axis of TANDEM is aligned with the 2-fold axis between the base pairs (Fig. 4).

Binding experiments with DNAs of varying base composition and with synthetic polydeoxynucleotides have revealed that TANDEM binds at least 7,500 times more tightly to poly(dA-dT) than to poly(dG-dC) and there are marked differences in the binding constants for DNAs of different nucleotide composition[7]. This specificity might be explainable in terms of hydrogen bonding between the peptide backbone of the drug and the DNA bases, using the model derived from the X-ray structure.

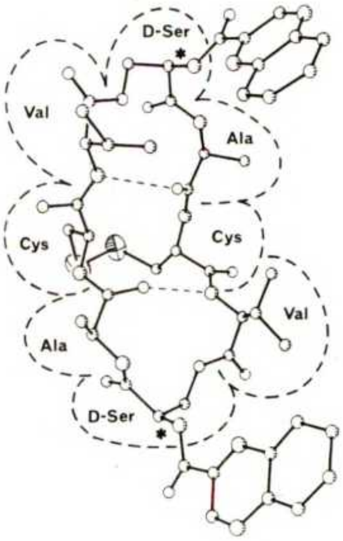


Fig. 2 Molecular structure of TANDEM determined by X-ray analysis at room temperature to a resolution of 0.85 Å. Dotted lines indicate intramolecular hydrogen bonds between the NH groups of L-Val and the O atoms of L-Ala. These bonds hold the molecule in a relatively stable conformation. Asterisks indicate the N–C$_\alpha$ (D-Ser) bonds through which the chromophores are attached to the octapeptide ring.

There are two potential hydrogen-bond donors and one potential acceptor in each half of the molecule (Fig. 3) which could be directed towards the interior of the helix on intercalation. The donors are the NH groups of L-Ala and D-Ser and the acceptors are the carbonyl O atoms of L-Ala. The latter are already involved in an intramolecular H bond with the NH group of L-Val (Fig. 2) and in an unfavourable position to accept a second hydrogen bond from the available donors of the DNA base pairs. Of the potential donor groups of TANDEM the NH groups of L-Ala seem to be well placed, as planes parallel to the chromophores oriented as in Fig. 3 and passing through these N atoms are separated by a perpendicular distance of 3.5–4 Å. In the intercalated complex, therefore, the NH groups of L-Ala lie co-planar with the sandwiched base pairs. Experiments with CPK models indicate that a reasonable fit can be obtained with hydrogen bonds of ~3 Å between the NH groups of L-Ala and the O-2 atoms of thymine in the narrow groove (Fig. 4). In this intercalated model an ApT sequence is sandwiched between the chromophores. The unwinding angle, spread over four base pairs, is 40°–50°, consistent with the experimentally determined value for DNA[2,7]. Attempts to build alternative intercalated models involving the N-3 atoms of adenine or the NH groups of D-Ser in hydrogen bonding were unsuccessful, nor could we achieve satisfactory intercalation via the major groove. It is known that binding of the related antibiotic, echinomycin, to phage T2DNA is not prevented by the presence of bulky sugar substituents which occlude the wide groove of the helix[4].

The X-ray structure of TANDEM prompts some suggestions about the character of the intercalation process. Interestingly, the isolated chromophore, quinoxaline-2-carboxamide, unlike simple monofunctional intercalators, shows no detectable interaction with DNA and does not affect the supercoiling of circular DNA[4,10]. Thus, the driving force for intercalation cannot be ascribed simply to a tendency to form stacking interactions between the chromophores and base pairs. In TANDEM, as shown by the crystal structure, the chromophores are held at the approximate position required for bifunctional intercalation by the relatively rigid depsipeptide backbone, and thus a very important geometrical constraint is already 'locked into' the structure of the ligand molecule. Moreover, all but two of the carbonyl O atoms are directed away from the chromophore side of the drug, where they are best placed to retain their hydrogen-bonding interactions with the surrounding solvent after binding to DNA. Such interactions are apparent in the present crystal structure where all but two of the carbonyl oxygens make at least one hydrogen bond with the surrounding water molecules and there is no direct hydrogen bonding between neighbouring TANDEM molecules. The hydration spheres are even more perfect in a structure of TANDEM determined at −135 °C (D. van der Helm, personal communication).

The naturally occurring antibiotics echinomycin and triostin A (Fig. 1*a*, *b*) display patterns of selectivity in binding to DNA and synthetic polydeoxynucleotides which, while characteristic for each antibiotic, are quite different from the marked preference of TANDEM for A + T-rich sequences[2,6,7]. For instance, echinomycin binds well to poly(dG-dC), approximately twice as strongly as it binds to poly(dA-dT), whereas the converse is true of triostin A. These differences from TANDEM are probably

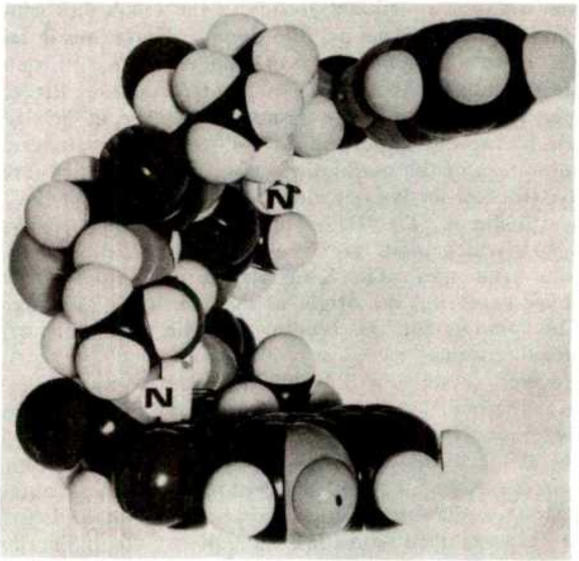


Fig. 3 CPK model of the TANDEM molecule with the quinoxaline chromophores rotated about the N–C$_\alpha$ (D-Ser) bonds from the crystal position to allow for bifunctional intercalation. The rotation is about 20° for each chromophore. The nitrogen atoms of L-Ala donating hydrogen bonds to the O-2 atoms of the thymine bases are indicated.

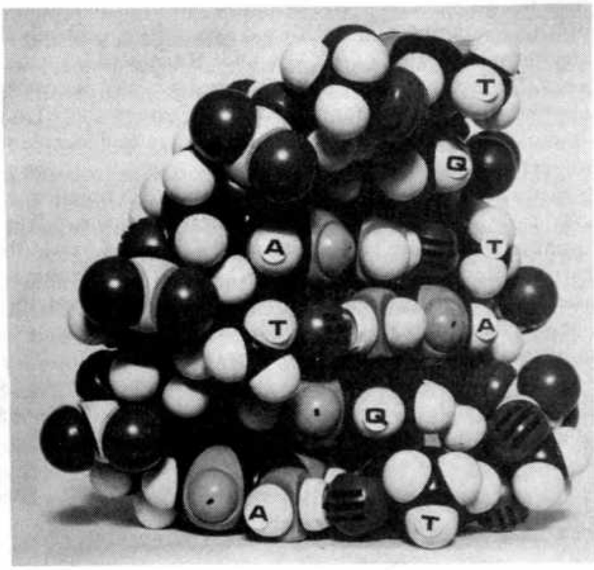


Fig. 4 CPK model of the intercalated complex between TANDEM and a partially unwound B-DNA helix showing the major groove side. The quinoxaline groups (Q) and the bases (adenine, A, and thymine, T) are labelled.

related to the effects of *N*-methylation on the precise conformation of the octapeptide ring and thus the availability of its functional groups for hydrogen bonding[2,4,5,13]. *N*-methylation of the L-Cys and L-Val residues in the natural antibiotics precludes formation of the intramolecular H bonds observed in TANDEM. Moreover, methylation of the L-Val residues necessarily results in conformational differences which will cause the carbonyl oxygens of L-Ala to project in the same direction as the chromophores. From this position they become available as effective hydrogen-bond acceptors for interaction with donors in the minor groove, that is, the 2-amino groups of guanine nucleotides, which may explain the ability of the antibiotics to bind to sequences containing G.C as well as A.T base pairs.

Extrapolation from the crystal structure of TANDEM to detailed molecular models for quinoxaline antibiotic–DNA complexes must be approached very cautiously. Although various lines of evidence (NMR[5,13], conformational calculations[5,11]) suggest that the peptide backbone in all these compounds adopts a relatively rigid and well defined conformation, there is no experimental evidence that this conformation does not alter on intercalation. Indeed, the significant differences found between the present room-temperature X-ray study and the similar one at −135 °C indicate that the orientation of some bonds, particularly those involving carbonyl groups, can be significantly affected by hydrogen bonding to solvent molecules[14]. There is also a degree of uncertainty about local DNA conformation[15,16] and about the orientation of the base pairs with respect to the helix axis in solution[17–19]. The recent X-ray study of the daunomycin–hexadeoxynucleotide complex[20] has shown that widespread changes can occur in the conformation of the backbone and the positioning of the bases on intercalation. Thus, although the present study gives well defined information about TANDEM itself, only a generalized model can be put forward for the DNA-drug complex. We are continuing our attempts to co-crystallize TANDEM with synthetic A-T deoxyoligonucleotides in the hope of establishing the structure of an intercalated complex directly by X-ray analysis.

We thank the MRC and the Cancer Research Campaign for financial support, the SRC for the diffractometer, the Deutsche Forschungsgemeinschaft for a postdoctoral fellowship (E.E.) and Professor R. A. Raphael for constructive comments. M.A.V. thanks the Department of Science and Technology, India, for financial support.

Received 5 August; accepted 24 December 1980.

1. Katagiri, K., Yoshida, T. & Sato, K. in *Antibiotics* Vol. 3 (eds Corcoran, J. S. & Hahn, F. E.) 234–251 (Springer, Heidelberg, 1975).
2. Waring, M. J. in *Antibiotics* Vol. 5, Pt 2 (ed. Hahn, F. E.) 173–194 (Springer, Heidelberg, 1979).
3. Waring, M. J. & Wakelin, L. P. G. *Nature* **252,** 653–657 (1974).
4. Wakelin, L. P. G. & Waring, M. J. *Biochem. J.* **157,** 721–740 (1976).
5. Cheung, H. T. *et al. J. Am. chem. Soc.* **100,** 46–54 (1978).
6. Lee, J. S. & Waring, M. J. *Biochem. J.* **173,** 115–128 (1978).
7. Lee, J. S. & Waring, M. J. *Biochem. J.* **173,** 129–144 (1978).
8. Sheldrick, G. M. (in preparation).
9. Fuller, W. & Waring, M. J. *Ber. Bunsenges. phys. Chem.* **68,** 805–808 (1964).
10. Waring, M. J. in *Drug Action at the Molecular Level* (ed. Roberts, G. C. K.) 167–189 (Macmillan, London, 1977).
11. Ughetto, G. & Waring, M. J. *Molec. Pharmac.* **13,** 579–584 (1977).
12. Wang, J. C. *J. molec. Biol.* **89,** 783–801 (1974).
13. Kalman, J. R., Blake, T. J., Williams, D. H., Feeney, J. & Roberts, G. C. K. *JCS Perkin* **I,** 1313–1321 (1979).
14. van der Helm, D. *et al.* (in preparation).
15. Wang, A. H. J. *et al. Nature* **282,** 680–686 (1979).
16. Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. *Nature* **286,** 567–573 (1980).
17. Levitt, M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 640–644 (1978).
18. Hogan, M., Dattagupta, N. & Crothers, D. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 195–199 (1978).
19. Hogan, M., Dattagupta, N. & Crothers, D. M. *Biochemistry* **18,** 280–288 (1979).
20. Quigley, G. J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).

Transglutaminase modifies the carboxy-terminal intracellular region of HLA-A and -B antigens

Jordan S. Pober* & Jack L. Strominger

The Biological Laboratories, Harvard University, Cambridge,
Massachusetts 02138, USA

The HLA-A and -B antigens, encoded in the major histocompatibility complex, are transmembrane proteins whose carboxy termini project inwards from the cell membrane. The intracellular region of the HLA-B7 antigen is known to include two glutamine residues[1]. We have used guinea pig transglutaminase to show that *in vitro* the enzyme couples amines specifically to glutamine residues in the carboxy-terminal region of the HLA-B7 and -A2 antigens. The specificity of this reaction has an immediate practical significance, but may also have a wider biological interest.

HLA-A and -B antigens contain an invariant extracellular polypeptide subunit (β_2-microglobulin or p12) of molecular weight (MW) 12,000 noncovalently attached to a polymorphic 44,000-MW transmembrane, glycosylated, heavy chain (p44)[2]. The heavy chain has three major regions: a 34,000-MW glycosylated, extracellular amino-terminal segment (p34) that is released from the membrane by papain in a complex with the light chain (p34, 12), a transmembrane hydrophobic segment of about 25 amino acid residues, and an intracellular hydrophilic carboxy-terminal region of about 30 amino acid residues[3]. Interactions of the HLA-A and -B antigens with intracellular

---

* Present address: Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA.

0028-0836/81/090819—03$01.00

constituents must involve the carboxy-terminal intracellular domain.

Transglutaminases are ubiquitous enzymes that catalyse the exchange of the γ-carboxamide ammonia group of a protein glutamine residue within a flexible region of polypeptide for an exogenous primary amine[4]. As shown in Fig. 1, guinea pig liver transglutaminase catalyses the insertion of ^{3}H-putrescine into the heavy chain of HLA-B7 antigen. Similar results were found with HLA-A2 antigen. When dansyl cadaverine was substituted for putrescine, the HLA-B7 or -A2 heavy chain became fluorescently labelled. Neither amine was incorporated into β_2-microglobulin (the HLA light chain). The location of the probe in the HLA heavy chain was ascertained by limited proteolysis. Papain cleaves native HLA-B7 antigen in detergent solution in a stepwise fashion, releasing first the carboxy-terminal hydrophilic region and then the penultimate hydrophobic region, altering the apparent MW of the heavy chain from 44,000 to 39,000 to 34,000 (ref. 3). As shown in Fig. 2, ^{3}H-putrescine radioactivity is released concomitantly with release of the intracellular carboxy-terminal domain. Thus the glutamine residue(s) that is modified must be one or both of those located within this region. Note that none of the 16 glutamine residues in the extracellular region seems to be modified. Therefore, only the intracellular hydrophilic region contains exposed glutamine residues in a flexible portion of the molecule.

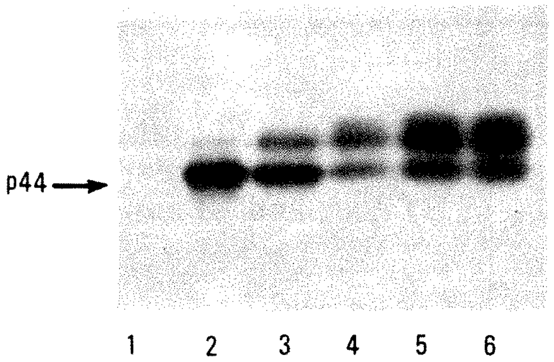


Fig. 1 Transglutaminase modification of HLA-B7. HLA-B7 antigen was purified from the membranes of JY lymphoblastoid cell membranes (doubly homozygous HLA-A2, -B7) by use of allospecific monoclonal antibody immunoaffinity chromatography in a modification (to be published elsewhere) of the procedure of Parham[9]. The pure antigen was adjusted to a concentration of 250 μg protein ml^{-1} in 10 mM Tris-HCl pH 8.0, 0.5% Nonidet P40 (NP40) non-ionic detergent. Guinea pig liver transglutaminase[10] at 50 μg ml^{-1} was pre-equilibrated in a solution made 500 mM in Tris-HCl pH 8.0, 0.5 mM putrescine, 40 mM CaCl$_2$ and 10 mM dithiothreitol at 37 °C for 30 min. This enzyme mixture (0.5 ml) and 0.5 ml of ^{3}H-putrescine-HCl in H$_2$O (0.5 mCi, NEN) were added to 1 ml of the HLA-B7 solution and the reaction was allowed to proceed at 37 °C. At different time points, 100-μl aliquots were withdrawn, placed on ice and made 25 mM in EDTA to inhibit the enzyme. Trichloroacetic acid (50 μl of 50%) was added to precipitate the protein. The precipitates were collected by centrifugation 1 h after the last time point was taken and were washed once with acetone to remove excess trichloroacetic acid. Each sample was dissolved in SDS sample buffer, subjected to SDS-polyacrylamide gel electrophoresis on a 20-cm 7–15% linear gradient with the buffer system of Laemmli[11], and analysed by Coomassie blue staining (not shown) and by fluorography[12]. Lanes 1–6 show the fluorograph of ^{3}H-putrescine radioactivity in the position of the HLA-B7 heavy chain after 30 s, 15, 30, 60, 120 and 180 min of transglutaminase-catalysed modification, respectively. Note that as labelling continues, a portion of the heavy chain shifts its electrophoretic mobility to one of slightly higher molecular weight. The explanation for this shift with time of incubation is unclear.

These results have two additional implications. First, exogenous transglutaminase may be used to modify specifically the carboxy-terminal region of HLA-A and -B antigens for various purposes. For example, spectroscopic probes[5] may be inserted to study the structure and function of this region of the HLA-A and -B molecules. HLA-A and -B antigens may be fluorescently or radioactively labelled, in a manner that does not perturb the antigenicity of the extracellular portion, for use in functional or binding assays with receptors for these antigens. Transglutaminase may be used as a vectorial label[6] to assess whether the carboxy-terminal region of the HLA-A and -B antigen is exposed in a reconstituted liposome preparation. Second, these data are consistent with the possibility that intrinsic transglutaminases could be involved in membrane events, including receptor-mediated uptake[7] and response to mitogenic lectins[8], by catalysis of the cross-linking of the intracellular portion of a transmembrane protein to other intracellular constituents. Labelling of HLA-A and -B antigens (or any other membrane proteins) in intact or broken peripheral blood lymphocytes or lymphoblastoid cells incubated with ^{3}H-putrescine or dansyl cadaverine has not been detected (J. P. and B. Guild, unpublished observations), but suitable conditions may not have been used. However, the data presented here establish that one

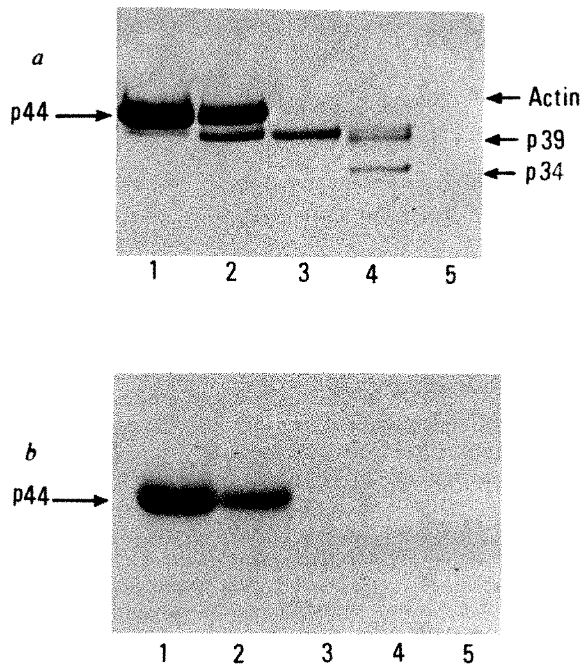


Fig. 2 Limited proteolysis of ^{3}H-putrescine-labelled HLA-B7. HLA-B7 antigen, labelled for 180 min as in Fig. 1 by transglutaminase-catalysed insertion of putrescine into the HLA heavy chain, was subjected to limited proteolysis by papain (Worthington). To 100-μl aliquots of HLA-B7 solution (at 125 μg ml^{-1}), in 100 mM Tris-H Cl pH 8.0, 25 mM EDTA, 10 mM CaCl$_2$, 2.5 mM dithiothreitol and 10 mM putrescine, 10 μl of papain were added and the solution was incubated for 1 h at 37 °C. The reaction was stopped by transfer to ice and by addition of 10 mM iodoacetic acid. The samples were trichloroacetic acid-precipitated and prepared for electrophoresis as in Fig. 1 legend. In this experiment an 8-cm 12.5% acrylamide gel was used with the buffer system of Laemmli[11]. Lanes 1–4 represent proteolysis by 0, 0.2, 1 and 5 μg of papain, respectively. Lane 5 is a rabbit skeletal muscle actin standard. *a* Shows the gel stained with Coomassie blue and *b* is a fluorograph of the same gel. Note that ^{3}H-putrescine label was excised concomitantly with release of the intracellular, hydrophilic portion of the HLA-B7 heavy chain as p44 was converted to p39.

common membrane protein in its intracellular domain is a suitable substrate for transglutaminase-catalysed reactions.

We thank Dr Lubert Stryer for guinea pig liver transglutaminase. This work was supported by research grant AI-10736 from the NIH and by a fellowship to J. S. P. from the NIH.

Received 19 November, accepted 31 December 1980.

1 Robb, R. J., Terhorst, C. & Strominger, J. L. *J. biol. Chem.* **253**, 5319-5324 (1978)
2. Springer, T. A., Strominger, J. L. & Mann, D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1539–1543 (1974).
3. Springer, T. A. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2481–2485 (1976)
4 Folk, J. E. & Finlayson, J. S. *Adv. Protein Chem.* **31**, 1–133 (1977)
5 Pober, J. S., Iwanij, V., Reich, E. & Stryer, L. *Biochemistry* **17**, 2163–2169 (1978)
6 Iwanij, V. *Eur. J. Biochem.* **80**, 359–368 (1977)
7 Levitzki, A., Willingham, M. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2706–2710 (1980)
8 Novogrodsky, A., Quittner, S., Rubin, A. L. & Stenzel, K. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1157–1161 (1978)
9 Parham, P. *J. biol. Chem.* **254**, 8709–8712 (1979).
10 Connellan, J. M., Chung, S. I., Whetzel, N. K., Bradley, L. M. & Folk, J. E. *J. biol. Chem.* **246**, 1093–1098 (1971)
11 Laemmli, U. K. *Nature* **227**, 680–685 (1970)
12. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–86 (1974).

(1→3)-β-D-Glucan (callose) is a probable intermediate in biosynthesis of cellulose of cotton fibres

H. Meier, L. Buchs, A. J. Buchala & T. Homewood

Institut de Biologie végétale et de Phytochimie,
Université de Fribourg, CH 1700 Fribourg, Switzerland

The biosynthetic pathway of cellulose in higher plants is unknown[1]. Attempts to produce cellulose *in vitro* have shown that mainly (1→3)-β-D-glucan (callose) is produced from particulate enzyme fractions of higher plants and uridine diphosphate glucose (UDPG)[2–5]. Some authors have claimed that in certain conditions, (1→4)-β-D-glucosidic linkages were also formed[6–10]. However, there is no conclusive evidence for the formation of pure (1→4)-β-D-glucans from UDPG. As callose forms naturally not only in various parts of plants (sieve and pollen tubes, root hairs, cell plates) but also in response to wounding, it seems best to study cellulose biosynthesis *in vivo* rather than in wounded cells or cell homogenates which might produce callose but not cellulose. Using intact cotton fibres, we have now shown by pulse-chase experiments that callose has a high turnover and is likely to be an intermediate in cellulose biosynthesis at the stage of secondary wall formation.

Naturally grown cotton fibres contain large quantities of callose[11–13]. In the fibres of one ovule of *Gossypium arboreum*, at the onset of secondary wall formation (about 25 days after anthesis), ~1 mg of callose is present out of a total fibre wall weight of ~13 mg. Using two experimental systems, we have investigated the incorporation of ^{14}C-labelled compounds into the cell wall polysaccharides of intact seed hairs of cotton (*Gossypium arboreum* L.). In the first system, whole fruit capsules received, through the cut surface of the petiole, a pulse of [U-^{14}C]sucrose (for about 20 min) followed by a chase with unlabelled sucrose for periods of up to 40 h (Fig. 1 legend). Six capsules, collected about 35 days after anthesis (at the stage of secondary wall formation), were treated in this way. At various intervals they were opened quickly and the intact seed clusters were immediately immersed in boiling 80% (v/v) methanol to avoid possible formation of wound callose. The fibres were isolated, further extracted with hot 80% methanol to remove low-molecular weight substances, and the methanol extracts combined; the fibres were finally homogenized. The radioactivity in the 80% methanol-soluble fraction was mainly contained in sucrose and may be taken to indicate the size of the low-molecular weight precursor pool at the time of analysis. The radioactivity in aliquots of the fibres was counted and other aliquots were treated with dimethyl sulphoxide (DMSO) to extract non-cellulosic high-molecular weight material[14]. The radioactivity in the latter and in the residues was measured. When the DMSO-soluble high-molecular weight material was hydrolysed, glucose was the chief product and the only labelled sugar obtained, although small amounts of unlabelled galactose, arabinose and xylose were detected. Methylation studies and radio gas chromatography, as well as degradation with exo-(1→3)-β-glucanase (EC 3.2.1.58) from *Basidiomycete* sp. QM 806, showed that the labelled glucan in the DMSO extract was of the (1→3)-β-type (callose-type) with a few (1→6)-linkages (our unpublished results). Further treatment of the residual fibre homogenate after DMSO extraction with the exo-(1→3)-β-glucanase gave only very minor amounts of glucose, indicating that with DMSO essentially all the (1→3)-β-D-glucan had been

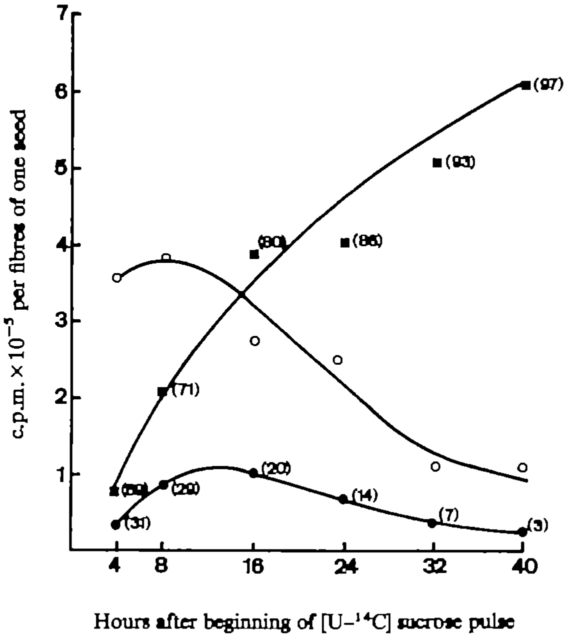


Fig. 1 Pulse-chase feeding of sucrose to petioles of fruit capsules of *Gossypium arboreum* L. The capsules were harvested 35 days post-anthesis (stage of secondary wall formation). The pulse of [U-^{14}C]sucrose (20 μl, 1 mM, 1.85 MBq) was applied through the cut ends of the petioles for about 20 min. The chase was with unlabelled sucrose (~50 μl h^{-1}, 1 mM) in the dark at 30 °C. After different incubation periods the radioactivity was measured in the 80% methanol-soluble fraction (○), the DMSO-soluble fraction, that is, (1→3)-β-D-glucan (callose) (●), and the residue, (1→4)-β-D-glucan (cellulose) (■), of the fibres. Numbers in parentheses are the percent radioactivity in callose- and cellulose-type glucans, respectively.

removed. No degradation was detected on treatment with lichenase (from *Bacillus subtilis*), α- or β-amylase, thereby excluding the possibility of the presence of other non-cellulosic glucans known to be present in higher plants. Acetolysis of the residue gave the peracetate derivatives of glucose, cellobiose and higher cellooligosaccharides. The cellobiose octaacetate was obtained in about 20% yield and such a value is usually taken to indicate that the polysaccharide in question is cellulose. This product was characterized by melting point, mixed melting point and TLC, and its specific activity, which was the same as that of the fibre residue, remained constant on recrystallization. These results show that the radioactivity of the fibre residue is essentially confined to cellulose.

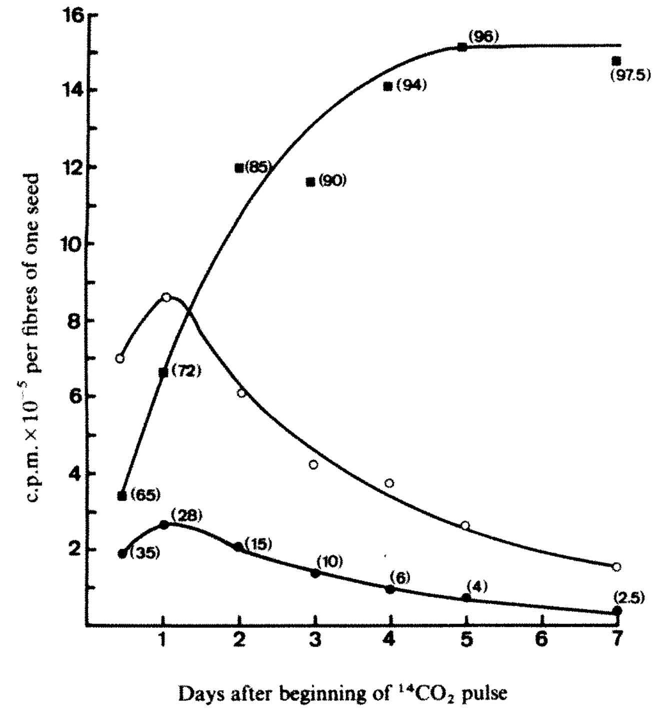


Fig. 2 Pulse-chase feeding of $^{14}CO_2$ to branches of intact plants of *Gossypium arboreum* L. $^{14}CO_2$ (7.4 MBq) was fed in the morning in bright sunlight to branches, each with three or four leaves and one capsule at 35–40 days post-anthesis, inside polythene bags. After 30 min about 80% of the original radioactivity in the bags had been taken up by the branches which were, however, left inside the bags for another 90 min. The latter were then removed and the plants were left in normal day and night conditions, when photosynthesis could occur normally, until the capsules were harvested. The radioactivity was determined in the 80% methanol-soluble fractions (○), the (1→3)-β-D-glucan (callose) (●) and the (1→4)-β-D-glucan (cellulose) (■) fractions of the fibres. Numbers in parenthesis are the percent radioactivity in the callose- and cellulose-type glucans, respectively.

Figure 1 shows that 4 h after the beginning of the pulse a high proportion (31%) of the radioactivity in cell wall glucans is confined to callose. The total radioactivity in the callose increased for about 16 h and then decreased continuously whereas the radioactivity in the cellulose continued to increase. A turnover of callose is evident; its radioactive content decreased from 100,000 c.p.m. in the fibres of one seed to 20,000 c.p.m. between 16 and 40 h after the pulse, while in the same period the radioactivity in the cellulose increased from about 400,000 to 600,000 c.p.m. This larger increase in the radioactive content of the cellulose is readily explained by the fact that 16 h after the pulse, there were still 280,000 c.p.m. in the fibres of one seed confined to low-molecular weight precursors (methanol extract). The apparent turnover of callose is relatively slow but this may be due to the slow replacement of radioactive sucrose by the unlabelled sucrose during the chase, or to the time delay in the arrival of sucrose to different fibres on different seeds in the capsule.

In the second system, $^{14}CO_2$ was given for about 2 h to branches of intact cotton plants with one capsule per branch. This seems to be the most natural experimental system in which abnormal wound reactions would be excluded. In this system also (Fig. 2), within a relatively short time after the pulse a high proportion (35%) of the 80% methanol-insoluble radioactivity of the fibres was in callose. The absolute radioactive content of the callose increased until about 1 day after the pulse and then decreased, whereas that of the cellulose continued to increase. The radioactivity in callose relative to that in cellulose decreased throughout the experiment. The radioactivity in the 80% methanol-soluble low-molecular weight substances of the fibres reached its maximum about 1 day after the beginning of the pulse and decreased thereafter. These results also indicate a turnover of callose and the probable incorporation of its glucose residues into cellulose.

A sharp chase with the nonradioactive compound was impossible with either of the two systems because of the time lag necessary for the entry of the chase substance into the fibre pool and because a certain amount of radioactive precursor continued to enter the fibre precursor pool from other parts of the fruit capsule. It is therefore feasible that the apparent rate of decrease in the radioactivity of the callose during the chase is lower than the apparent rate of increase in the radioactivity of the cellulose even if callose is an obligatory precursor for cellulose. Although our results, in contrast to those of Maltby *et al.*[15], indicate a clear turnover of callose, they do not show whether or not callose is a direct glucosyl donor for cellulose. If this is the case, a (1→3)-β-D-glucan:(1→4)-β-D-glucan glucosyl transferase (transglucosylase) would probably be involved in cellulose synthesis. (1→3)-β-D-Glucanases are undoubtedly present in cotton fibres (unpublished results) and we are now investigating whether, in certain conditions, they can act as transglucosylases.

Suggestions that callose is a cellulose precursor are not new and have been made by several plant cytologists[16–18]. Waterkeyn (personal communication) showed by fluorescence microscopy, after staining with aniline blue, that at different stages of secondary wall formation, cotton fibres always contain a thin layer of a substance reacting like callose at the inner side of the wall. If further studies confirm the precursor role of callose for cellulose in cotton fibres as well as in other cell types, it could be argued that callose may accumulate where the biosynthetic step from callose to cellulose is inhibited.

This work was supported by the Swiss NSF.

Received 28 August; accepted 4 December 1980.

1. Delmer, D. P. *Recent Adv. Phytochem.* **11**, 45–77 (1977).
2. Feingold, D. S., Neufeld, E. F. & Hassid, W. Z. *J. biol. Chem.* **233**, 783–788 (1958).
3. Ordin. L. & Hall, M. A. *Pl. Physiol.* **43**, 473–476 (1968).
4. Flowers, H. M., Batra, K. K., Kemp, J. & Hassid, W. Z. *Pl. Physiol.* **43**, 1703–1709 (1968).
5. Heiniger, U. & Delmer, D. P. *Pl. Physiol.* **59**, 719–723 (1977).
6. Villemez, C. L., Franz, G. & Hassid, W. Z. *Pl. Physiol.* **42**, 1219–1223 (1967).
7. Péaud-Lenoël, C. & Axelos, M. *FEBS Lett.* **8**, 224–228 (1970).
8. Smith, M. M. & Stone, B. A. *Biochim. biophys. Acta* **313**, 72–94 (1973).
9. Van der Woude, W. J., Lembi, C. A., Morré, D. J., Kindinger, J. I. & Ordin, L. *Pl. Physiol.* **54**, 333–340 (1974).
10. Raymond, Y., Fincher, G. B. & MacLachlan, G. A. *Pl. Physiol.* **61**, 938–942 (1978).
11. Meinert, M. C. & Delmer, D. P. *Pl. Physiol.* **59**, 1088–1097 (1977).
12. Huwyler, H. R., Franz, G. & Meier, H. *Pl. Sci. Lett.* **12**, 55–62 (1978).
13. Huwyler, H. R., Franz, G. & Meier, H. *Planta* **146**, 635–642 (1979).
14. Pillonel, Ch., Buchala, A. J. & Meier, H. *Planta* **149**, 306–312 (1980).
15. Maltby, D., Carpita, N. C., Montezinos, D., Kulow, K. & Delmer, D. P. *Pl. Physiol.* **63**, 1158–1164 (1979).
16. Eschrich, W. *Protoplasma* **47**, 487–530 (1956).
17. Currier, H. B. *Am. J. Bot.* **44**, 478–488 (1957).
18. Geuns-Longly, B. & Waterkeyn, L. *C.r. hebd. Séanc. Acad. Sci., Paris* D**283**, 761–763 (1976).

MATTERS ARISING

Allez Neanderthal

APSIMON[1] concluded that Lévêque and Vandermeersch's discovery of a Neanderthal specimen at Saint-Césaire associated with Châtelperronian denies a Neanderthal ancestry for modern Europeans. In reality, taking account of other evidence a completely different conclusion may be drawn from this critical discovery.

The following points are relevant. (1) Robust but modern (although not European) populations are associated with Mousterian (Middle Palaeolithic) industries in the Levant[2,3]. (2) Neanderthal (or Neanderthal-like) remains are associated with Aurignacian (Upper Palaeolithic) in Croatia[4–6]. (3) Evolutionary trends in the European Neanderthals themselves (*contra*[7]) are clearly in the direction of modern Europeans[6,8]. (4) The earliest modern Europeans (earlier than Cro-Magnon) are distinctly Neanderthal-like in their morphology[5,8–10].

The evolutionary hypothesis is supported by the backward extension of trends in the modern Europeans (the samples become distinctly more Neanderthal-like earlier in time) and the forward extension of trends in the Neanderthals of Europe (they become distinctly more like modern Europeans later in time). In this regard, Saint-Césaire specifically resembles the late (more modern) Neanderthals from Vindija cave[4–6] and thus provides important support for the observed trends within the Neanderthal sample.

The *average* changes between late Neanderthals and early modern Europeans should not be confused with the comparison of La Chapelle and the Cro-Magnon male. Such changes could happen fairly quickly, responding to changes in skeletal stress during growth[11] and changes in selection[5]. If migration played an important role, the origin of the migrating populations with their distinctive European features remains unknown. If the Levant is considered to be the place of origin, in spite of the lack of morphological relations, the resulting juxtapositions of hominids and industries makes nonsense of virtually every previous hypothesis of archaeological or morphological sequence.

MILFORD H. WOLPOFF

Department of Anthropology,
University of Michigan,
Ann Arbor, Michigan 48109, USA

1. ApSimon, A. M. *Nature* **287,** 271–272 (1980).
2. McCown, T. D. & Keith, A. *The Stone Age of Mount Carmel* 2 (Clarendon, Oxford, 1939).
3. Vandermeersch, B. in *The Origin of Homo sapiens* 49 (ed. Bordes, F.) (UNESCO, Paris, 1972).
4. Malez, M. *et al. Curr. Anthrop.* **21,** 365 (1980).
5. Wolpoff, M. H. *Palaeoanthropology* (Knopf, New York, 1980).
6. Wolpoff, M. H. *et al. Am. J. phys. Anthrop.* (in the press).
7. Stringer, C. B. *J. archaeol. Sci.* **1,** 317 (1974).
8. Frayer, D. W. *Evolution of the Dentition in Upper Palaeolithic and Mesolithic Europe* (University of Kansas, Lawrence, 1978).
9. Szombathy, J. *Die Eiszeit* **2**(1), 73 (1925).
10. Schwalbe, G. *Studien zur Vorgeschichte des Menschen* (Schweizerbartsche, Stuttgart, 1906).
11. Brothwell, D. R. *J. archaeol. Sci.* **2,** 161 (1975).

APSIMON REPLIES TO WOLPOFF—We are in agreement about point (1). Concerning point (2): The relevant hominid finds (from Vindija, layer G_1) are described as 'non-diagnostic morphologically' and could be Neanderthal or early *sapiens*, the lithic artefacts from the layer are not diagnostic, but the single split-based bone point is consistent with 'Aurignacian'. There is, however, no intrinsic improbability about Neanderthal—Aurignacian association. I agree with points (3) and (4). Analyses which bring together in one sample all classic European Neanderthal crania and in another sample all European Upper Palaeolithic crania (including some specimens now considered to be Mesolithic) obscure this. The recently described Neanderthal frontal bone from Hahnöfersand near Hamburg[1], dated to ~36,000 BP, which has been considered as indicating hybridization with *sapiens*, might equally well be cited as a further example of this evolutionary trend.

In general, the hypotheses presented by Wolpoff provide an explanation of the Neanderthal-*sapiens* transition in Europe which can be far more easily harmonized with current interpretation of the archaeological evidence than can the hypotheses presented by Stringer *et al.* However, if the archaeologist must pick his way with care through the differing interpretations offered by anthropologists, the latter would be well advised to remember that 'Upper' and 'Middle Palaeolithic', 'Mousterian', 'Aurignacian' and such are artificial constructions, that attempts to attribute artefactual assemblages to one or other category may create impressions of discontinuity where none exists, and that the most useful information to be obtained from artefacts relates to patterns of activity.

1. Bräuer, G. *Archäol. Korresp.* **10,** 1 (1980).

WE are in broad agreement with ApSimon[1] regarding the significance of the Saint-Césaire discovery from France[2], although it must be emphasized that our comments are preliminary in the absence of detailed descriptions of the archaeology and fossil human material. An effect of this discovery will be to focus the 'Neanderthal problem' in Europe to a consideration of events in the period between ~35,000 and 30,000 yr b.p.

The discovery of early Upper Palaeolithic tools in association with a Neanderthal skeleton at Saint-Césaire has apparently confirmed the prediction that the technological transition to the Châtelperronian was accomplished by indigenous populations[3]. It is now plausible that the same kind of association between Neanderthal hominids and industries termed 'early Upper Palaeolithic' applies for other areas in Europe (for example, for the 'leaf-point' industries of northern and central Europe and for the Uluzzian of Italy)[4]. Furthermore, the cultural continuity of these industries (for example, the Châtelperronian) with later Upper Palaeolithic industries (the later Perigordian [Gravettian] for example) has by no means been established because consideration of the subsequent development of these early Upper Palaeolithic industries often reveals the intervening presence of Aurignacian assemblages which are typologically not locally derived. In France there are occurrences of both Châtelperronian and Aurignacian, but during the subsequent Arcy interstadial (~32,000–31,000 yr b.p.) the Châtelperronian disappears[5,6].

ApSimon's claim that there is no reliable evidence for the existence of the Aurignacian complex in central Europe earlier than it occurs in France needs qualification. Although the Tŭrnovo province of Bulgaria is not geographically central in Europe, recent work there at the Bacho Kiro cave is highly germane to a discussion concerning the origin of the Aurignacian in Europe. For this site there is now an internally consistent series of radiocarbon dates which supports the existence of a well defined Aurignacian industry at about 43,000 yr b.p. (Ly-1102, GrN-7569, 7545, 7570) which has no local antecedents. It is claimed that the same stratum is associated with 'primitive' but anatomically modern human remains[7,8]. If we accept these dates from Bacho Kiro it raises the possibility that the Aurignacian there is ancestral to that of central and western Europe. However a postulated east–west diffusion of Aurignacian populations and/or techniques cannot yet be demonstrated with the present precision of the radiocarbon method.

In south-west Asia the critical time period available for an evolutionary transformation or replacement of Neanderthal populations is probably before 40,000 yr b.p., perhaps well before[9]. Thus *the Saint-Césaire find apparently* demonstrates for the first time that

0028-0836/81/090823—02$01.00

Neanderthals in western Europe probably existed at the same date as much more modern-looking populations in south-west Asia (as at Skhūl and Qafzeh) and perhaps in Europe (as at Velika Pećina, Brno, Mladeč (*contra* ApSimon) and Hahnöfersand)[10–16]. No well preserved fossils which are demonstrably intermediate in significant characters between Neanderthals and anatomically modern humans have been described from either of these areas. The fact that some of the earliest anatomically modern hominids in these areas display characters such as robust brow-bridges and large teeth does not necessarily imply that they evolved from Neanderthal ancestors, as such characters are probably merely plesiomorphous (primitive) for *Homo sapiens* generally, and therefore have no phylogenetic relevance in these comparisons. As European and south-west Asian Neanderthals share suites of characters which are apparently autapomorphous (uniquely derived), it is these traits which are significant in any analysis of relationship between Neanderthal and early anatomically modern hominids, as is the presence or absence of traits considered autapomorphous for anatomically modern hominids. With very few exceptions such characters do not cross the morphological division between the Neanderthal and early anatomically modern groups[10–13,17].

The many characters shared between the European and south-west Asian Neanderthals and the many different characters shared between the early anatomically modern fossils of the same two areas mitigates against any evolutionary scheme which posits only *in situ* evolution between Neanderthals and anatomically modern humans in each area[10–13,17]. Evidence now suggests that Eurasian Neanderthals were temperate or cold-adapted forms, whereas the early anatomically modern humans of the same area were not. This implies that a model involving at least some gene flow from outside this area at the time of the first appearance of anatomically modern humans is more appropriate[18].

Thus we agree with ApSimon that, pending further description, the Saint-Césaire Neanderthal disproves the idea of an *in situ* hominid transformation in Europe at a Mousterian/Upper Palaeolithic interface. Instead, the polyphyletic origins of the Upper Palaeolithic in Europe may be matched by a dichotomy, at least, in the morphology of the hominids that produced it. The wealth of new evidence that is now accumulating, together with the promise of a more reliable chronological framework, will allow us to monitor more precisely the events of this key period in prehistory. We would also argue that more sophisticated models are needed than those which posit either mass extinction, hybridization or rapid *in situ* evolution of Neanderthal hominids at the time of the appearance of anatomically modern humans in Europe.

C. B. STRINGER
R. G. KRUSZYNSKI
Department of Palaeontology,
British Museum (Natural History),
Cromwell Road, London SW7 5BD, UK

R. M. JACOBI
Department of Classics and Archaeology,
Lonsdale College,
University of Lancaster, Bailrigg,
Lancaster LA1 4YN, UK

1. ApSimon, A. M. *Nature* **287,** 271–272 (1980).
2. Lévêque, F. & Vandermeersch, B. *C.r. hébd. Séanc. Acad. Sci., Paris* D **291,** 187 (1980).
3. Bordes, F. in *The Origin of Homo Sapiens* (ed. Bordes, F.) 211–218 (UNESCO, Paris, 1972).
4. Bánesz, L. & Kozłowski, J. K. (eds) *L'Aurignacien et le Gravettien (Périgordien) dans leur Cadre Écologique.* (Institut Archéologique de l'Académie Slovaque des Sciences à Nitra, 1980).
5. Delporte, H. *Bull. Antiq. Nat.* **1,** 15–28 (1969).
6. Leroi-Gourhan, A. & Renault-Miskovsky, J. in *Approache Écologique de l'Homme Fossile* (eds Laville, H. & Renault-Miskovsky, J.) 35–49 (Université Pierre et Marie Curie, Paris, 1977).
7. Kozłowski, J. K. *Przegl. Archaeol.* **28,** 77 (1979).
8. Delporte, H. & Djindjian, F. *Przegl. Archaeol.* **28,** 101 (1979).
9. Farrand, W. R. *J. archaeol. Sci.* **6,** 369 (1979).
10. Vandermeersch, B. thesis, Univ. Paris VI (1977).
11. Stringer, C. B. *J. archaeol. Sci.* **1,** 330 (1974).
12. Stringer, C. B. in *Recent Advances in Primatology. Vol. 3 Evolution* (eds Chivers, D. J. & Joysey, K. A.) 414 (Academic, London, 1978).
13. Trinkaus, E. & Howells, W. W. *Scient. Am.* **241,** 118 (1979).
14. Smith, F. H. *Am. J. phys. Anthrop.* **44,** 127 (1976).
15. Jelínek, J. *Anthropologie* **14,** 79 (1976).
16. Bräuer, G. *Z. Morph. Anthrop.* **71,** 1 (1980).
17. Stringer, C. B. & Trinkaus, E. in *Aspects of Human Evolution* (ed. Stringer, C. B.) (Taylor & Francis, London, in the press).
18. Trinkaus, E. in *Aspects of Human Evolution* (ed. Stringer, C. B.) (Taylor & Francis, London, in the press).

APSIMON REPLIES TO STRINGER *ET AL.*—Concerning Bacho Kiro: If the industry corresponds to accepted definitions of 'Aurignacian', 43,000 BP would be a surprising date, as it would be ~10,000 yr older than dates obtained elsewhere.

We can agree that Velika Pécina, Brno and Mladeč may be of the same order of age as Saint-Césaire, subject to the uncertainties of present chronologies, but not that they might be substantially older, as was implied by the traditional interpretation of Würm I-II-III in central Europe.

I do not believe that the case has yet been made for the previous evolution of the Aurignacian. Interstratification of Aurignacian and Châtelperronian raises the possibility that they are functionally differentiated variants of the same tradition. Blade technology is not substantially more important in these early European 'Upper Palaeolithic' industries than in some Mousterian industries and development from the Mousterian lithic tradition seems to be their most likely origin.

ARTHUR APSIMON
Department of Archaeology,
University of Southampton,
Southampton SO9 5NH, UK

Why do jackals help their parents?

BLACK-BACKED jackals (*Canis mesomelas*) sometimes stay with their parents and help them raise subsequent litters. Based on a positive correlation between the number of pups surviving in a litter and the number of adults (parents plus helpers) in the family unit, Moehlman[1] concluded that pup survivorship is enhanced by the presence of helpers. Unfortunately, her analysis contains errors of calculation and interpretation that seriously weaken this conclusion. Here I correct those errors and present an alternative explanation for her results.

First, an error in the calculation of the correlation coefficient (between 'pups surviving' and 'number of helpers') markedly inflates the statistical significance of the result. From the data shown in Moehlman's Fig. 1, I calculate $R_s = 0.58$ (Spearman rank correlation coefficient; $P < 0.05$) and not $R_s = 0.967$ ($P < 0.01$) as reported. Although the correlation coefficient is still statistically significant, factors other than the number of jackal helpers in a family unit obviously contribute substantially to the observed variation in the number of pups that are raised from a litter.

Second, it is stated[1] that "an adult helper gains more (yield 1 pup per adult) by being a helper of its parents than by finding a mate and raising its own pups aided only by its mate (yield $\frac{1}{2}$ pup per adult)." For both parents and helpers, however, the relevant measure of reproductive success is the total number of pups that survive from a litter, and not the number that survive per adult. As long as a helper's parents stay mated, the coefficients of relatedness between the helper and its siblings in subsequent litters, and between the parents and their new offspring, are the same (that is, $r = \frac{1}{2}$). Therefore, a helper realizes the same 'reproductive success' whether helping its parents increase their litter size by n pups or raising n pups itself (with its own mate). The calculation of 'yield per adult' always underestimates the benefit that accrues to helpers when they contribute to the increased reproductive success of their parents. For example, in terms of yield per adult, a jackal seems to increase its reproductive success by only $n/3$ when helping its parents raise an additional n pups, but it increases its reproductive success by $n/2$ by raising its own n pups. In fact, both of these increases in reproductive success should be the same.

Data shown in Moehlman's Fig. 1 can now be used to assess the reproductive consequences of becoming a helper instead of mating. The slope of the geometric mean regression[2] of 'pups surviving' over 'number of helpers' ($y = 0.9 + 1.67x$; $n = 15$) is an estimate of the

0028-0836/81/090824—02$01.00

average increase in reproductive success of a family unit due to the addition of each helper. Thus, each helper adds about 1.7 pups to a litter raised by its family unit. The reproductive success of an adult that finds a mate and raises pups, without any helpers, is, on average, 1.0 pups ($n = 4$). In the short term, then, jackals do seem to gain more (a net gain of $1.7 - 1.0 = 0.7$ pups) by helping their parents than by mating, but these gains must ultimately be weighed against the disadvantages of foregoing both dispersal and mate selection until the family unit breaks up (for example, when one of the parents dies). It is not surprising, therefore, that few jackals remain with their parents for more than 1 yr (ref. 1).

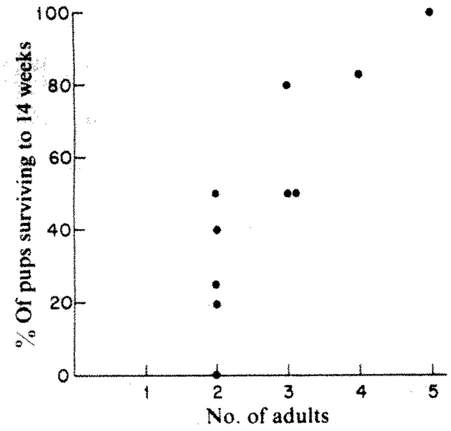


Fig. 1 Number of adults per family and pup survivorship from 3 to 14 weeks ($n = 10$; Spearman rank correlation coefficient, $R_s = 0.88$; $P < 0.01$).

Finally, Moehlman erroneously equates 'number of pups surviving' with 'pup survivorship'. Survivorship means survival rate, and would be an estimate of the average probability of a pup surviving to weaning, but such data are not presented. The litter size of black-backed jackals at parturition varies from one to eight pups[3]. I suggest, therefore, that the positive correlation between the number of helpers and the number of pups raised may not be caused by pup survival rates, but may be the result of higher litter sizes at birth for females that have helpers. Helpers are usually young jackals that have stayed with the family unit from a previous litter[1]. Their continued presence could, therefore, act as a proximate stimulus for their mother to increase litter size, especially as helpers also provide food for the family unit[1]. This pattern should be looked for in other studies of communal breeding systems and may be expected in any species where clutch size varies with food supply. For example, such a response to helpers seems to occur in both the Tasmanian native hen (*Tribonyx mortierii*)[4] and the Puerto Rican Tody (*Todus mexicanus*)[5], where clutch sizes are larger when families have helpers. Jackals may ultimately help their parents because helping increases their inclusive fitness but further work will be needed to determine whether those fitness increases are due to larger litter sizes or increased pup survival.

ROBERT D. MONTGOMERIE

Department of Biology,
Queen's University,
Kingston, Ontario, Canada K7L 3N6

1. Moehlman, P. D. *Nature* **277**, 382–383 (1979).
2. Ricker, W. E. *J. Fish Res. Bd Can.* **30**, 409–434 (1973).
3. Brand, D. J. *Proc. zool. Soc. Lond.* **140**, 617–659 (1963).
4. Ridpath, M. G. *C.S.I.R.O. Wildl. Res.* **17**, 1–118 (1972).
5. Kepler, A. K. *Publ. Nuttall Orn. Club* **16**, 1–190 (1977).

MOEHLMAN REPLIES—Montgomerie is correct in recalculating $R_s = 0.58$ (Spearman rank correlation coefficient; $P < 0.05$). I thank him for his correction; however, I disagree with his alternative explanation, in that he implies it is mutually exclusive.

The correlation coefficient is still significant. Furthermore, litter no. 15 was an aberrant group in which the parents were old[1], only one helper of three was functional throughout the critical period of pup development and one pup survived. This particular helper was also the only example of an individual helping for more than 1 yr. This datum could be shifted to the three-adult column, but for biological reasons I think that it distorts the relationship between presence of helpers and pup survival. If correlation coefficients are calculated for the remaining 14 litters, $R_s = 0.83$ ($P < 0.01$).

The relevant measure of reproductive success is the number of genes surviving. Thus Montgomerie is correct in stating that the important measure is the total number of surviving, related pups. Additional data for three litters were collected in 1978, and on average, in 17 litters: one pair raised 1.3 pups ($n = 6$); one pair + 1 helper raised 3.3 pups ($n = 8$); one pair + 2 helpers raised 4.0 pups ($n = 2$); one pair + 3 helpers raised 6.0 pups ($n = 1$). (The data on litter 15 are not included in these calculations.) Thus, by delaying reproduction for 1 yr, a helper potentially increases its inclusive fitness by on average 1.7 pups, assuming that helping incurred a minimal cost. At present there is controversy as to how to measure 'inclusive fitness' and any suggestions would be appreciated. This, of course, is only one variable, and factors such as increased experience on the natal territory, territorial inheritance and increased survivorship may be equally important (from the helper's point of view). The situation is further complicated by the fact that a 1-year-old jackal can emigrate and still accrue this inclusive fitness as long as a sibling stays and helps. Data on success at emigrating, mating and establishing a territory at 1 yr of age compared with 2 yr of age are needed to answer fully questions concerning kin selection in this study population of black-backed jackals.

Number of adults versus number of live pups at 14 weeks is the critical measure in terms of kin selection and investigating the problem of why jackals help their parents. However, to clarify the relationship between the presence of helpers and pup survival, further data are presented (Fig. 1). Litter size at first emergence from the den (approximately 3 weeks old) is known for 10 of 17 litters. Dens were not disturbed because of the potential detrimental effect on pup survival and subsequent den use (dens may be used again). The mother spends most (90%) of her time in the den during the first 3 weeks, and is fed by her mate and helpers. Data presented in ref. 2 was for pup survival to 14 weeks. As stated, it was during this vulnerable period that mortality was observed. Thereafter, pups no longer used a den and began to forage throughout the home territory. Number of adults versus pup survivorship (3–14 weeks): $n = 10$, $R_s = 0.88$, $P = < 0.01$; number of adults versus pup survival at 14 weeks: $n = 17$, $R_s = 0.85$, $P = < 0.01$; number of adults versus litter size at 3 weeks: $n = 10$, $R_s = 0.83$, $P = < 0.01$; number of pups at 3 versus 14 weeks: $n = 10$, $R_s = 0.68$, $P = < 0.05$.

The highest R_s value is between number of adults and percentage of pups surviving from 3 to 14 weeks. The lowest R_s value is between litter size at 3 weeks and 14 weeks. Information on gestation and prenatal growth in mammals suggests that pup size rather than number of pups will be affected by maternal nutrition, except in cases of prolonged starvation.[3] Litters of dogs raised in the laboratory experience high mortality during the first 3 weeks[1]. Thus information on litter size at 3 weeks may not be a reliable indicator of initial litter size. Helpers may increase initial litter size, but in current data obtained with jackals, the most significant correlation is with pup survivorship from 3 to 14 weeks. Thus it seems reasonable to conclude that helpers improve pup survival rates, possibly from a very early age, and may even affect initial litter size.

PATRICIA D. MOEHLMAN

Department of Zoology,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Anderson, A. C. *J. Am. vet. med. Ass.* **130**, 151–158 (1957).
2. Moehlman, P. D. *Nature* **277**, 382–383 (1979).
3. Hendrickx, A. G. & Houston, M. L. in *Reproduction and Breeding Techniques for Laboratory Animals* (ed. Hafez, E. S. E.) 157–176 (Lea & Febiger, Philadelphia, 1970).

Size constancy does not fail below half a degree

AS an observer's distance from an object varies, so does the visual angle subtended by the object at the observer's eye. One of the most reliable phenomena in psychology is that when distance information is available the perceived size of an object is nearly constant despite variations in visual angle, or retinal image size[1]. A corollary is that images of constant visual angle seem to vary in size with perceived distance (Emmert's law). This phenomenon ('size constancy') is only one part of the more general property of perception that one might call 'object constancy', embracing constancy of size, shape, depth, velocity, lightness, colour and so on. A recent report[2] has claimed "that size constancy holds when image size is above half a degree, and breaks down when it is less". The claim, however, was based on demonstrations and anecdotal evidence, rather than experiment. We show experimentally that size constancy in the laboratory is equally good at target sizes above and below $\frac{1}{2}$°.

The targets were single, black, vertical lines 1.2 arc min wide presented at three viewing distances (114, 228 and 456 cm) with four lengths ($\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, 1 deg). Each line was fixed in the centre of a white card subtending 5.25° vertically by 3.75° horizontally independent of viewing distance. Each card was placed at the appropriate position on a long laboratory bench on which papers and small objects were scattered. Viewing was binocular with full room lighting. Thus, distance information was abundant, but for a given target size (angle), the length and width of the line and of the frame in which it was set were of constant visual angle, independent of distance. The subject looked back and forth between the target and an oscilloscope screen placed at 228 cm distance, about 6° to the right of the target. His task was to adjust the vertical separation between two small green dots on the screen until the separation 'appeared to match the size of the target'. Trials were presented in random order and the 12 conditions were repeated 3 times for each of the 10 subjects (5 male, 5 female).

The results (Fig. 1) show clearly that perceived length of a line of constant visual angle varies with distance in almost exactly the same way for all target sizes in the range $\frac{1}{8}$°–1°. Analysis of variance not surprisingly showed highly significant effects of both visual angle and viewing distance ($P \ll 0.001$), but most importantly the interaction between distance and visual angle was not significant ($F_{6,29} = 1.47$, $P > 0.10$). If size constancy were weaker or absent below half a degree the slope of the curves in Fig. 1 would be shallower or even flat for small target lengths. Our data strongly contradict any such effect. Constancy was not complete (this would be given by a slope of 1) but it was equally good for all target sizes.

In a second part of the experiment, an electronic flashgun was used to create the afterimage of a line $\frac{1}{4}$° long. The observer viewed the afterimage on a plain white card of constant visual angle at each of the three distances, and made length matches as in the first part of the experiment. A second or third exposure of the line was given if necessary to complete a total of nine trials. One subject reported no afterimage, but the data from the remaining nine (crosses in Fig. 1) were not significantly different from those obtained with a real line of $\frac{1}{4}$° ($F_{1,8} = 0.84$, not significant).

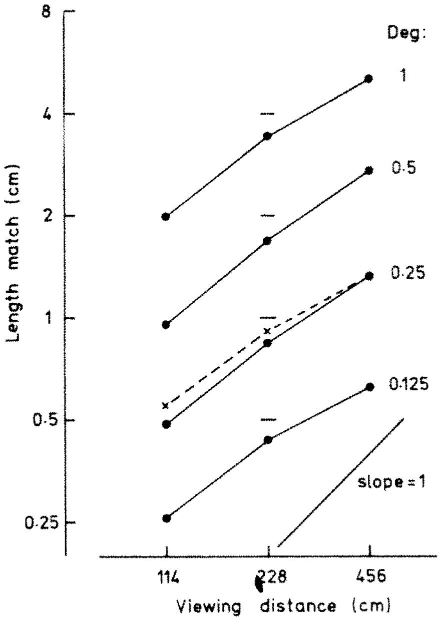


Fig. 1 Solid symbols show results averaged over 10 subjects who adjusted a pair of dots to match the lengths of lines subtending four different visual angles at three viewing distances. Crosses represent similar judgements (nine subjects) made with an afterimage $\frac{1}{4}$° long. Horizontal bars represent the ideal match when target and dots were at the same distance. A slope of 1 would represent perfect constancy, and zero slope would indicate judgements based purely on visual angle (no constancy). The average slope of regression lines fitted to the solid points was 0.70, with no significant interaction between distance and visual angle.

Ross *et al.*[2] reported that afterimages subtending less than $\frac{1}{2}$° did not seem to change size with 'projected' distance, and that the finding was robust in classroom demonstrations. We flashed a 2-cm line to a group of 27 students, all of whom were more than 228 cm distant. The afterimage would thus be less than $\frac{1}{2}$° for all observers. They observed the image on their hands and then on a far wall. None reported no change of size, 19 said the image got bigger when further away, 2 said it got smaller and 6 reported no image. Ross *et al.*[2] also appeal to the reader's direct experience of the breakdown of constancy. Our own experience confirms our classroom demonstration, that small afterimages obey Emmert's law in the usual way. When we acted as subjects in the experiment, our data were virtually identical to those of Fig. 1.

Ross *et al.*[2] remark that objects seen in the distance or from the top of a tall building seem 'small'. It may be that in these cases, and in their Fig. 1, information about the true distance is inadequate, or impoverished by poor resolution of detail. As a result, both distance and size would be underestimated[1]. The fact that many distant objects subtend less than $\frac{1}{2}$° may be quite incidental.

We conclude from our experiments and demonstrations that size constancy for both real targets and afterimages is just as good for small images as for large ones. Indeed, we should suffer very curious distortions in the perceived relative size of objects if this were not so. There seems to be no substance to the claim that size constancy fails below half a degree.

We thank the SRC for financial support.

M. A. GEORGESON
M. G. HARRIS

University of Bristol,
Department of Psychology,
8-10 Berkeley Square,
Bristol BS8 1HH, UK

1. Rock, I. *An Introduction to Perception*, Ch. 2 (Macmillan, New York, 1975).
2. Ross, J., Jenkins, B. & Johnstone, J. R. *Nature* **283**, 473 (1980).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

0028-0836/81/090826—01$01.00

BOOK REVIEWS

Risk assessed and risk perceived

Eric Ashby

Societal Risk Assessment: How Safe is Safe Enough? Edited by R.C. Schwing and W.A. Albers Jr. Pp.363. (Plenum: 1980.) $29.50, £18.59.

LOOKED at with the naive eye of logic, a calamity killing 10^4 people, likely to occur once every 10^4 years, does not appear different from a calamity killing one person every year. Looked at with the complex eye of human judgement, there is a massive difference. A nuclear melt-down once every ten years, killing half a million people, is a horrifying thought; but that level of carnage is only the linear extrapolation of the average annual death rate on the roads of the United States. This disparity between the statistical assessment of a risk and the personal perception of the risk has important implications for public policy. The cost of making things safe — roads, airlines, nuclear power stations, drugs — rises steeply as the risks are reduced. A level of cost is reached when it would be fantastically expensive to make the thing any safer. So politicians have to ask: How safe is safe enough?

It is not surprising, therefore, that conferences and books on risk-management are cropping up all over the place. They bring together the opinions of statisticians, who can tell you with some confidence what your chances are of being killed on the roads, stifled with emphysema through smoking cigarettes, or struck by lightning; and the opinions of psychologists, who can tell you (with as much confidence, if with less evidence) how you and I evaluate these statistical assessments. And the conferences are attended by policy-makers who go away perplexed and under the obligation to make decisions which will reconcile the disparity between assessed and perceived risks. The disparity remains, but it is encouraging that conferences on risk-management bring to the surface new ideas and fresh attitudes to a problem which has become of major political importance.

Societal Risk Assessment is a record of one of the best recent conferences on the issue. It was arranged by the General Motors Research Laboratories. Let it be said at once that although General Motors might be assumed to have a vested interest in reassuring the public about some kinds of risks, there is no sign of bias in this volume. There are frank discussions of the risks we run and the risks we accept; this section includes an excellent summary of the causes of road accidents, studied in England by a team at the Transport and Road Research Laboratory. The book then goes on to discuss what it costs to reduce risks. Notable among these discussions is a paper from an expert on underwriting in the insurance business. When you come to think of it, insurance underwriters are the one profession who have to try to reconcile assessed and perceived risks; otherwise they would go out of business.

Having discussed what are, and what are not, acceptable risks, and how much society can afford to pay to reduce risks, the conference then went on to tackle an even more controversial question: Who shall decide what risks we are permitted to run? This is a touchy issue among motorists in Britain just now. Despite the alarming casualty figures on the roads of Britain (nearly 7,000 a year killed) the chance of any individual in the whole population being killed is only 1 in 2,500 years. This explains the indifference of the British public toward road accidents. But the problem for society as a whole remains serious: every fatal accident is reckoned to cost some £64,000. It is not disputed that the number of fatal and serious accidents would be reduced if people wore seat belts, drove at not more than 55 miles an hour (as they are obliged to do in many American States), and did not drink alcohol before driving. Is it a duty of the State to compel its citizens to take these precautions? And if so, can the citizens be persuaded by propaganda to come to a consensus about this (after all, there is a consensus about driving on the left-hand side of the road)? Or must they be compelled without having consented? Or, in the name of liberty, should motorists be allowed to endanger themselves (at the taxpayer's expense in a country with a national health service) to their heart's content? The conference did not, of course, come to decisions on these questions; but the issues were thoroughly aired.

The first three sections of the book endorse, with a wealth of fresh examples and neat arguments, conclusions already familiar to those who have studied the subject: for example that there is no such thing as zero risk; that people (including experts) commonly bias their opinions by giving too much weight to the severity of an incident and too little weight to its frequency; that unfamiliar risks, such as hazards from radioactivity and chlorofluorocarbons, generate more anxiety than "classical" risks, such as floods and tornadoes. The last section of the book tackles the consequence of these conclusions, namely that "at the center of the risk problem are people and their fears". Life in affluent Western countries is more secure than it has ever been; yet people are more jittery about hazards than their grandparents ever were. There are still highly educated people who will not sit down to dinner with 13 at the table, and other highly educated people who join expeditions to climb in the Himalayas. The last three papers in the book deal with this fascinating region of social anthropology.

There is already a considerable library of books on risk assessment. Two of the books are outstanding in quality and content. One is the report on the international seminar held in Stockholm in 1978 (*Energy Risk Management*, edited by G.T. Goodman and W.D. Rowe; Academic Press, 1979). The other is *Societal Risk Assessment*. Anyone who has digested these two books can reckon to be up to date in the state of the art of risk management. ☐

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Electron microscope? Better buy the book!

R. K. S. Wood

Fungus Diseases of Tropical Crops. By Paul Holliday. Pp.607. (Cambridge University Press: 1980.) £55, $125.

A LARGE proportion of the world's crops are grown in the tropics where they are of two main kinds. First are those with cash products, important in world trade and grown in large units using advanced technologies. Then there are the crops, not always of food, grown in much smaller units by methods still largely traditional and little affected by modern technology. As would be expected, there is substantial research on diseases of the first type of crop, but much less on diseases of the other, although, of course, these diseases are as or more important when they decrease the yields of subsistence foods to tragically low levels. Diseases of food crops in temperate countries may be financially inconvenient; in the tropics they can cause famine.

There are many general texts and

innumerable monographs on diseases of temperate crops — understandably, because most of the world's plant pathologists are in non-tropical countries. In relation to need there are all too few authoritative texts on diseases of tropical crops; the publication of a new and large book on the subject is, therefore, something of an event.

Dr Paul Holliday has been studying and reporting on diseases of tropical crops for about 30 years, many of them as a member of a pool of plant pathologists, spending much of his time in the tropics and the rest of it at the Commonwealth Mycological Institute at Kew. With this background, there was every reason to expect a first-class book from the author, and first class it is.

An introduction describes in detail the plan of the book and its parts, and the main sources of information on taxonomy and nomenclature. There follow, in 556 double-column pages of small but readable type, descriptions of fungal pathogens and the diseases they cause in plants grown as crops in the tropics. The fungi are listed alphabetically by perfect state genera with, occasionally, the omission of entries under the imperfect state genera better known to many plant pathologists. A description of the morphology and more general features of each genus is followed by a list of main references to the genus. For each species, under each genus there is a description of the morphology and a critical reference followed by accounts of the diseases it causes — again accompanied by a list of important references, each including the volume and page or abstract number for an abstract in the *Review of Applied Mycology* or the *Review of Plant Pathology;* this is particularly useful. The accounts are full and remarkably compact for the amount of information they contain. Quite properly, the emphasis is on factors relevant to diseases in the field and on their control. But there are quite adequate references to physiological, biochemical and related research.

There is an excellent appendix of host plants, their major and minor pathogens, and selected references, a list of common and botanical names of host plants and, finally, two indexes of fungi under perfect and conidial state names.

This book presents a huge amount of information accurately, clearly and systematically, is an excellent work of reference and will remain so for many years. The accounts of the diseases are such that they can also be read with pleasure. Expensive though it be, a thousand copies could be bought for the price of one electron microscope and, where most needed, they will be a lot more useful. They will also require little maintenance and last much longer. ☐

R. K. S. Wood is Professor of Plant Pathology and Head of the Department of Pure and Applied Biology at Imperial College, University of London.

Topical tropical meteorology

Reginald Newell

Climate and Weather in the Tropics. By H. Riehl. Pp.611. (Academic: 1979.) £28, $67.50.

STRANGERS to our subject may wonder why the tropics should be accorded a separate treatment in studies of climate and weather. After reading Riehl's book, I suspect that it is the sheer volume of new material available rather than fundamental differences in the characteristic phenomena that prompts this division. Many of the early meteorological studies dealt with the tropics; Trade Winds were analysed for their practical use to shipping, and early ideas about the global general circulation were based (often erroneously) on the tropical Hadley cell circulation.

In the twentieth century, frontal systems and the changing patterns of cyclones and anticyclones began to be explored in depth, both observationally and theoretically, and the focus and excitement switched to middle latitudes. The old ideas of general circulation were replaced by those which gave a pre-eminent role to the large-scale quasi-horizontal eddy processes; theoretical work by Jeffreys was followed by observational studies by Bjerknes, Priestley, and Starr and his colleagues. Most of the observational data came from the middle and high latitudes of the Northern Hemisphere. During and after the Second World War, forecasters from middle latitudes had to ply their trade in the tropics and many tried to translate ideas about middle-latitude systems to those of the tropics, without success. In fact, when data on the general circulation of the tropics became available it was realized that the quasi-steady Hadley cell circulation and the associated Walker circulations in longitudinal planes dominate the tropics, while the eddy processes govern the general circulation of middle latitudes.

There are of course some major differences between the regions: the rather flat surfaces of constant potential temperature in the tropics compared to the steep slopes — the strong baroclinicity — of middle latitudes; the easterly jet streams rather than the westerly jets of middle latitudes; the presence of a two-year cycle, with a shift from east to west, in the zonal winds of the tropical lower stratosphere as compared with the one-year seasonal cycle of middle and high latitudes; the large extent of tropical warm ocean, with temperatures greater than about 300 K, which seems to be necessary to spawn and maintain hurricanes; the mean vertical motions which in monsoon regions reverse from winter to summer rather than every few days as in middle latitudes; the dominance of the Earth's vorticity over the local vorticity in middle latitudes; and the strong vertical coupling which characterizes systems of middle latitudes and contrasts with the near independence of the lower and upper troposphere in the tropics. In this book, Riehl describes all the important phenomena and places them in the context of the global general circulation. He seems to have read all that has been written on tropical meteorology and provides an excellent set of references. A notable feature is the inclusion of findings from recent expeditions to the tropics such as the Tropical Atlantic Experiment.

The problems of how energy is transported vertically in the tropics from its source, effectively at the sea surface, and how energy gets into the atmosphere from the warm ocean — both topics that Riehl has contributed to significantly — are dealt with extensively and there is ample material for students to carry the subjects further.

Tropical cyclones and the various ideas about their structure, mechanisms, formation and movement are covered in detail and occupy about one quarter of the book. There is a well-written chapter by Ferdinand Bauer on numerical prediction of hurricanes.

My only major complaint — from which Bauer's chapter is exempt — is that numerous typographical errors are scattered throughout the book; together with the awkward construction of some phrases and poor standardization of units, they often make the text hard to follow. A firm editorial hand needs to be applied before the next edition. Nevertheless, Riehl's book provides a unique source of information and deserves to be read by all meteorologists and to be the standard reference for many years. ☐

Reginald Newell is a Professor of Meteorology in the Department of Meteorology and Physical Oceanography at the Massachusetts Institute of Technology.

tRNA 2

J.P. Goddard

Transfer RNA: Biological Aspects. Edited by D. Söll, J.N. Abelson and P.R. Schimmel. Pp.578. (Cold Spring Harbor Laboratory: 1980.) $70, $84 outside USA.

THE publication of this book follows approximately six months after that of its companion volume *Transfer RNA: Structure, Properties and Recognition* (reviewed in *Nature* **284**, 287; 1980). The format of the two volumes is similar. Thus the contributors, who were participants at the Transfer RNA Meeting at Cold Spring Harbor in August 1978, have given us here a series of readable, authoritative and objective articles covering all the biological aspects of tRNA, except synthetase recognition, aminoacylation and ribosome interaction which were dealt with in the first volume. Division of the contributions into four sections — biosynthesis, gene arrangement, suppression and coding, and "other roles of tRNA" — is perhaps the best in the circumstances. However, the organization and expression of tRNA genes are so intimately linked that articles in the first two sections overlap and some of them also impinge upon suppression which has proved such a useful property in study of tRNA transcription and processing.

The opening section on tRNA biosynthesis begins, as do most sections in both volumes, with a useful general review of the topic by the session chairman at the meeting. There then follow a dozen shorter articles in which specific genetic and enzymological studies are thoroughly covered. Gene arrangement is introduced by a beautifully succinct review. The topics of the articles that follow, though broad-ranging, are somewhat unbalanced with four articles on *Drosophila*, including three on *in situ* hybridization, but none on tRNA genes in mitochondria or in *Xenopus*. No general reviews introduce the final two sections of the book. While a brief overview of the rather esoteric topic of suppression would have been most useful, the diverse topics of the final section could only be dealt with satisfactorily by individual articles. The volume concludes with three appendices on tRNA gene localization and codon usage.

One final comment should be made on the two-year delay in publication which, in some respects, is fortunate rather than regrettable. In mid-1978 there were tantalizing promises of rapid advance in some of the newer fields of study, such as tRNA gene organization in eukaryotes and removal of intervening sequences from tRNA precursors. The contributors have made an admirable job of updating their articles — of the 1300 references cited, nearly 200 date from 1978, 75 from 1979 and 16 from 1980. Thus most of the discoveries in the year following the 1978 meeting are reported in this volume.

The book has all the qualities one has come to expect from Cold Spring Harbor monographs and should find an appreciative readership among workers in tRNA research and in molecular biology generally. ☐

J.P. Goddard is a Lecturer in Biochemistry at the University of Glasgow.

Epithelial physiology: tribute to Jean Maetz

Clive Ellory

Epithelial Transport in the Lower Vertebrates. Edited by B. Lahlou. Pp.366. (Cambridge University Press: 1980.) £25, $55.

THIS book represents a memorial symposium held as a tribute to Jean Maetz by many of his friends, colleagues and collaborators. It was a meeting Jean would certainly have enjoyed, the emphasis being on fish and amphibian epithelial physiology, with many authors concentrating on endocrinological aspects. Aside from the obvious environmental attractions of working in Villefranche, Naples or Woods Hole, physiologists have become increasingly interested in fishes for studying the mechanisms of epithelial secretion. Fish gills and opercular skin, flounder intestine and shark rectal gland are all preparations which have generated valuable models and influenced thinking on epithelial function. These systems exercised Jean Maetz's considerable talents and represent the principal content of the present volume.

Papers on fish gills comprise half the total contributions. Thus, gill structure and function are covered well in the initial section, with emphasis on vasculature and functional morphology. The importance of the chloride cell in gill physiology is apparent from the number of times this topic recurs in the book, and from the descriptions of the exciting, isolated skin preparations from *Fundulus* and *Gillichthys*. New results, including data on gill permeability to large molecules, and some detailed electron microscopy produce tentative functional models for the chloride cell. Other gill transport systems beside Cl covered in the book include Na^+/NH_4^+ exchange and calcium transport. There are good biochemical presentations on gill Na, K-ATPase and the more controversial anion-stimulated ATPase. A salutary lesson is presented in the Na,K-ATPase chapter, where vanadate, a specific inhibitor of this enzyme *in vitro*, was found to be functioning on gills *in vivo* as a vasoconstrictor.

Kidney is mostly approached from an endocrinological point of view, although there are two impressive studies on single nephron physiology which emphasize the widely different roles for the kidney in marine and freshwater fish. Intestinal functions discussed include salt and amino acid transport, and amphibian skin is represented by three contributions mostly concerned with epithelial sodium transport. Endocrinological factors are emphasized in many of the contributions, but there are two small reviews dealing specifically with the hormonal control of ion transport.

Overall, there is much of interest in this book. Selection criteria for participants in such a memorial symposium obviously include factors which are not strictly scientific or rigorously thematic, which is reflected by some unevenness, with certain elements sitting less easily under the general title. There are also clear differences in scientific content, philosophy and merit among the various contributions, including some which are trivial or frankly eccentric. However, on balance this is a fitting tribute to Jean Maetz, which should prove valuable to fish physiologists and epithelial transport workers. ☐

Clive Ellory is a Lecturer in Physiology at the University of Cambridge.

Jean Maetz , a talented physiologist

Birds in brief

Robert Hudson

A Complete Checklist of the Birds of the World. By Richard Howard and Alick Moore. Pp.701. (Oxford University Press: 1980.) £17.50, $49.50.

THIS book comprises a straightforward listing of the living birds of the world, giving for each species the scientific and selected English name and a brief indication of geographical range; for polytypic species, named subspecies are also listed. Essentially, it is a condensed and partly updated version of the authoritative 15-volume *Checklist of the Birds of the World* edited by J.L. Peters and others (Harvard University Press, 1931–1979), but lacking the taxonomic and nomenclaturial comments, references and detailed distribution notes to be found in the larger work. The introduction to the present book suggests its major use will be as a handy personal checklist for the widely-travelled birdwatcher. Perhaps, but I will find it more useful as a quick reference for identifying unfamiliar bird names encountered in the serial literature, and in this respect its 59-page index is admirable.

Specialists in particular bird groups will not find it hard to quibble with points of detail, but on the whole the compilers offer a reasonable compromise between the often conflicting treatments to be found in avifaunal publications. Several other books of this type have appeared since 1974, the present one being the most comprehensive and therefore the most costly. It is the only one to include subspecies; but this is the weakest aspect of the book, since its authors are not taxonomists. A simplified listing of racial names gives a misleading impression of subspecific rigidity (whereas clinal situations are common even within land areas as small as the British Isles); moreover, named races present a continuum from well-differentiated semi-species to those hardly worthy of recognition, though this basic truth is not made evident to the reader. Thus, this is a potentially useful book, but not a definitive work. □

Robert Hudson is on the scientific staff of the British Trust for Ornithology, and is editor of the Trust's quarterly journal Bird Study.

Semiconductor lasers come of age

S. D. Smith

Physics of Semiconductor Laser Devices. By G. H. B. Thompson. Pp.549. (Wiley: 1980.) £22.75, $68.25.

THE semiconductor diode laser based upon the radiative recombination of carriers injected across a p–n junction is now nearly 20 years old from the date of the original concept. It has developed from a laboratory speciality operating only at cryogenic temperatures into a range of optoelectronic devices of application in optical fibre communication, military systems, printing, infrared spectroscopy, pollution detection and short-pulse work, to mention only part of the list.

This excellent and comprehensive book opens with a most readable chapter outlining the history, concepts and basic principles, without the use of mathematics, emphasizing the dramatic effect of the introduction of hetero-structures, using layers of alloy material, which separately confine the injected carriers and guide the generated optical waves with favourable effects on threshold current and temperature of operation. In the next 100 pages the author discusses the basic processes of light emission and laser action, and includes appropriate development of absorption, stimulated and spontaneous processes, band structure in the presence of heavy doping, and recombination times. Gain–current relationships, threshold currents and efficiencies are calculated and compared with experimental results for the basic or homostructure p–n junction laser with a simple resonator, for which the mode properties are also given. The treatment owes something to the earlier *Semiconductor Lasers and Light Emitting Diodes* by Kressel and Butler (Academic, 1977), but is more continuous and complete at the expense of rather intricate notation. I found this section a useful text for part of a graduate course on optoelectronic devices held at the Optical Sciences Center of the University of Arizona.

Later chapters deal with heterostructure lasers, optical waveguides and all the many variations of laser geometries in considerable detail. Methods of fabrication and the physics of dynamical response are also considered.

The only notable omission is the absence of discussion in any detail of the lead salt family of diode lasers and their application in high resolution infrared spectroscopy, perhaps not surprising since the book concentrates on the GaAs alloys. However, the basic physical principles for all types of semiconductor diode lasers are well covered which is a stated aim of the book. The author does not claim to be comprehensive in citing the literature — a difficult task in this active field — but nevertheless provides copious references in each chapter.

The book frequently examines the relation between semiconductor lasers and other types of lasers — emphasizing the high gain values, for example. It is thus easy to see what semiconductor injection lasers have not achieved — visible room temperature operation, high powers or wide-band tuning in one device.

With the slight qualification that the book's title is not quite consistent with its concentration on the GaAs injection diode laser family, it is to be generally recommended for a variety of uses in research, application and teaching. □

S. D. Smith is Visiting Professor at the Optical Sciences Center, University of Arizona, Tucson, and Head of the Physics Department at Heriot-Watt University, Edinburgh.

Organic infrared–structure correlations

P.C. Crofts

Infrared Characteristic Group Frequencies. By G. Socrates. Pp.153. (Wiley: 1980.) £24, $72.

INFRARED spectroscopy is perhaps the most generally applicable and widely used physical method for investigating structures of organic compounds. Although qualitatively rationalized by classical physical theory, it is an essentially empirical activity in which absorption bands reveal individual chemical bonds or larger submolecular units whose vibrations interact with radiation of the observed frequencies. Users of the technique thus require comprehensive collections of assignments of recorded absorption frequencies to particular functional groups or other structures. Although described in the preface as "a simple introduction to characteristic group frequencies", this book is a substantial and valuable work, simple only in its arrangement and empiricism and not in any restriction of its subject matter.

An introductory chapter contains three successively more detailed charts of the range 4000–200 cm^{-1}; the first shows regions without absorptions and infers absences of particular groups, the second relates observed bands to possible vibrations, and the third shows absorption regions for many specified compound types. This is followed by 19 chapters dealing with individual substituents of particular importance (e.g. O–H, C = O), with collections of related substituents (e.g. triple-bonded groups, halogen compounds), or with larger structural features (e.g. alkane residues, six-membered heterocycles). A short section surveys X–H stretching overtones in the 14000–4000 cm^{-1} region. Each chapter contains extensive tables of absorption bands, in wavenumber and wavelength units, with well-written explanatory text and many references.

A compilation of this type can be only a starting point in an infrared investigation — which also requires comparison with reference spectra of model compounds — but when used intelligently this book should point many investigators towards solution of their problems. It is well printed, with wide, short, double-column pages which are convenient when examining a spectrum, and is recommended to all who use infrared spectroscopy in structural investigations.

P.C. Crofts is a Senior Lecturer in Chemistry at the University of Manchester Institute of Science and Technology.

CORRESPONDENCE

Continued from page 742

anomalies in the data very successfully. For instance, he showed that in the Cornish sample the biggest single annual decline in the incidence of herd breakdowns was from 5.5 to 3.2 per cent in 1975–76 (a 42 per cent decline). Yet by 31 August 1976 only ten "fire-brigade" cases had been completed in Cornwall[1], an area of approximately 355,000 hectares (a fire-brigade action normally involves the removal of a small number of badger social groups around a farm with a herd breakdown). Is Dr Yates telling us that this 42 per cent decline in Cornwall was the instant result of ten completed fire-brigade actions? If ten fire-brigade actions could be so successful, it is surprising that the TB problem was not solved long ago. In other areas a similar situation can be seen, with significant reductions in the number of herd breakdowns before the onset of badger gassing, or before effective measures could have been achieved by the badger gassing programme. The data are not conclusive, but there is good evidence to suggest that the incidence of TB was undergoing a decline before the onset of gassing, and that to attribute the decline solely to the gassing campaign is not justified.

Dr Yates also contested that only 15 per cent of the Cornish herd breakdowns could be attributed to badgers, and "presumed" that badgers were responsible in many more cases. Lord Zuckerman (*Nature* 19 February 1981, p.628) also contested this point, saying that badgers in Cornwall have only been routinely investigated as a possible source of infection since 1974. However, his figures still demonstrate our original point; even since 1974 badgers were still only believed responsible for a minority of breakdowns, and even these figures are very dubious. In his letter Lord Zuckerman states that the majority of herd breakdowns in Cornwall supposed to have originated from badgers were attributed to infected badgers living up to two miles from the breakdown, and 28 per cent were attributed to infected badgers found more than two miles away. But in areas of relatively high badger density, such as Cornwall, badger territories are small, about 75 hectares (report, p.37). Ministry of Agriculture, Fisheries and Food (MAFF) studies in Gloucestershire[2] showed that in an area of high badger density, itinerant badgers were rare, and long distance movements outside the group territory were not recorded. The available information suggests that in Cornwall badgers are very unlikely to make contact with a cattle herd two miles away. Lord Zuckerman might as well tell us that in 100 per cent of cases an infected badger was found within twenty miles; such a figure is no less meaningful than those presented by Lord Zuckerman.

Figures such as those discussed are of little value if we do not understand how and where transmission takes place. As we said in our original letter, the mere presence of infected badgers does not always result in herd breakdowns. In his letter, Lord Zuckerman states that transmission of TB from badgers to cattle occurs in pasture that has been contaminated with sputum, pus, urine and faeces. It has never been proven that this is the sole point of transmission, nor even the major point of transmission, and this is another area in which further research is needed. As long ago as 1974, MAFF staff[3] pointed out that "On many occasions badgers have actually been observed in cattle buildings, presumably in search of food, and contamination of cattle food concentrates may also be important [in transmission]" and "On a number of occasions farmers have reported finding badgers in farm buildings eating cattle feeding stuffs". In his report, Lord Zuckerman (p.35) said that badly diseased badgers may leave their sett and take up refuge in farm buildings. How common is such behaviour? Further investigation might show that farm buildings are a major point of transmission, and if so a significant reduction in herd breakdowns might be achieved by a simple improvement in animal husbandry. Actions such as these to complement the gassing programme do not seem to have been considered by Lord Zuckerman; we feel that they should be.

Lord Zuckerman asserts that during the period of the moratorium (October 1979 to October 1980) the increase in the percentage of reactor herds in the affected counties of the South West was a direct result of halting badger control measures. This assumption was supported by Dr Yates in his letter, where he showed that the incidence of TB in badgers increased *immediately* following the moratorium on gassing. This not only totally ignores the likelihood of natural cyclic trends in TB prevalence, as discussed in our original letter and the reply by Dr Yates, but also requires us to believe that following the moratorium there was an immediate increase in the badger population, and consequently of the numbers of badgers infected with TB. It must be remembered that although no new areas were gassed during the moratorium, gassing was continued "to maintain freedom from the disease in areas already cleared of it"[2]. Since the moratorium only applied to areas not previously gassed, it seems reasonable to presume that, in the absence of any evidence to the contrary, the badger population was fairly stable in those areas. Badgers are long-lived animals with a relatively slow reproductive rate. Since the moratorium only lasted one breeding season, even if there was a population increase it could only have been small. And even presuming that there was an upward trend in badger population density during the period of the moratorium, there is no unequivocal evidence that TB in badgers is density dependent. Lord Zuckerman's report (p.68, 70) shows that TB is not only prevalent in badgers in areas of high density, but also in areas such as South Dartmoor, where the badger population is very low. Also, when considering Lord Zuckerman's and Dr Yates' arguments, it is necessary to take account of the rate of development of the disease in badgers. Experiments by MAFF scientists suggest that this may take several months[1]. In affected areas of the South West, tuberculin testing of cattle is carried out annually[3], and so there will be a further delay between possible infection of cattle and their subsequent discovery at tuberculin testing. Clearly there is no evidence to support Lord Zuckerman's and Dr Yates' argument that TB in badgers and cattle immediately increased as a result of the moratorium on gassing, and it is not surprising that an increase in TB in badgers was not seen in all areas, the point we made in our original letter.

Much of Lord Zuckerman's latest letter deals not with the TB problem, but is an attack on the Mammal Society and some of its members. Lord Zuckerman questions the right of the society to discuss the subject, since he suggests that the society is not fully competent in all the relevant fields of expertise. While this is a debatable point, no one has ever questioned Lord Zuckerman's right to express his opinion.

Also, Lord Zuckerman is unfamiliar with the aims or constitution of the Mammal Society, and so it is necessary to correct some more misleading statements. We are not a conservation society, and such interests are outside our remit. However, at our AGM on 14 April 1973, a motion from the membership mandated the Council:- (1) "to encourage the discussion of contentious issues [relating to the study of mammals] with the aim of showing where more information is needed in order to reduce the uncertainties of the issue", and (2) "to write, when appropriate, general letters asking for action or expressing concern". These are quotations from the minutes of that meeting, and clearly we have acted in accordance with the wishes of our membership. Lord Zuckerman's lengthy discussion on whether other societies can act in a similar manner is totally irrelevant. And of course our statements are not anonymous; they have been made by the officers of the Mammal Society, and a full list of elected officers is published.

The Mammal Society hopes that the badger/TB problem can be solved soon, and that some means can be found whereby the incidence of TB in cattle in the South West can be permanently reduced. We still feel that this aim can only be achieved by a critical analysis of all the data. We feel that the Zuckerman report does not help solve this complex problem; the only statistical treatment presented in the report is the calculation of percentage changes. This is clearly unacceptable for a complex epidemiological problem. Since the data are not conclusive, and since none of our original points have been answered satisfactorily, we suggest that the Minister of Agriculture should establish a Scientific Advisory Group, containing a small number of independent scientists such as wildlife epidemiologists and statisticians, who can analyse the data in detail and advise the Consultative Panel on Badgers and Tuberculosis, which currently lacks such expertise.

J. R. Flowerdew
*(Hon. Sec.,
Mammal Society)*

Reading, Berks, UK

1. Ministry of Agriculture, Fisheries and Food *Bovine Tuberculosis in Badgers* (HMSO, London, 1976).
2. Ministry of Agriculture, Fisheries and Food *Bovine Tuberculosis in Badgers: Fourth Report* (HMSO, London, 1980).
3. Muirhead, R.H., Gallagher, J. & Burn, K.J. *Vet. Rec.* 95, 552–555 (1974).

BOOKS RECEIVED

Chemistry

PAUL, J.K. (ed.). Large and Small Scale Ethyl Alcohol Manufacturing Processes from Agricultural Raw Materials. Pp.576. ISBN 0-8155-0815-8. (Noyes Data Corporation, Park Ridge, New Jersey: 1980.) $48.

SIVAKUMAR, T.C. Instrumental Methods of Analysis: A Laboratory Manual. Pp.118. (Chemical Engineering Education Development Centre, Indian Institute of Technology, Madras: 1980.) Pbk $10.

Technology

McAULIFFE, C. A. Hydrogen and Energy. Pp.109. ISBN 0-333-18432-7. (Macmillan, London: 1980.) £12.

Computer Science

McGETTRICK, A. D. The Definition of Programming Languages. Cambridge Computer Science Texts — 11. Pp.268. Hbk ISBN 0-521-22631-7; pbk ISBN 0-521-29585-8. (Cambridge University Press: 1980.) Hbk £13.50; pbk £6.25.

Earth Sciences

ALLEN, R.H. and STURDY, R.G. Soils in Essex III. Sheet TL71 (Little Waltham) Soil Survey Record No.62. Pp.172. (Soil Survey, Rothamsted Experimental Station, Harpenden, Herts: 1980.) Flexi £3 with 2 maps, £1 without.

AVERY, B.W. Soil Classification for England and Wales (Higher Categories). Technical Monograph No.14. Pp.67. (Soil Survey of England and Wales, Rothamsted Experimental Station, Harpenden, Herts: 1980.) Flexi £1.

BARTON, R. The Oceans. Pp.336. ISBN 0-87196-414-7. (Facts on File, New York: 1980.) $19.95.

BISWAS, A. K. (ed.). The Ozone Layer. Environmental Sciences and Applications, Vol.4. Pp.381. Flexi ISBN 0-08-022429-6. (United Nations Environment Programme/Pergamon: 1979.) Np.

BRANICA, M. and KONRAD, Z. (eds). Lead in the Marine Environment. Proceedings of the International Experts Discussion on Lead Occurrence, Fate and Pollution in the Marine Environment, Rovinj, Yugoslavia, October 1977. Pp.352. Flexi ISBN 0-08-022960-3. (Pergamon: 1980.) $60, £27.

DELYANNIS, A. and DELYANNIS, E.-E. Seawater and Desalting, Vol.1. Pp.189. ISBN 3-540-10206-X. (Springer-Verlag: 1980.) DM 59, $34.70.

DIX, H. M. Environmental Pollution: Atmosphere, Land, Water, and Noise. Pp.286. Flexi ISBN 0-471-27905-6. (Wiley: 1981.) £5.95, $17.50. Also available in hbk.

EMBREY, P. G. and FULLER, F. P. (eds). A Manual of New Mineral Names 1892--1978. Pp.467. ISBN 0-19-858501-2. (British Museum (Natural History)/Oxford University Press: 1980.) £24.

INSTITUTE OF APPLIED GEOLOGY. Evolution and Mineralization of the Arabian–Nubian Shield, Vols 1,2,3 and 4. Proceedings of a Symposium convened by Ahmad M. Al-Shanti, Jeddah, Saudi Arabia, February 1978. Vol.1, Pp.200. ISBN 0-08-024660-2; Vol.2, Pp.224. ISBN 0-08-024467-X; Vol.3, Pp.120. ISBN 0-08-024480-7; Vol.4, Pp.168. ISBN 0-08-024481-5. (Pergamon: 1980.) Vols 1, 2 and 3 $30, £15; Vol.4. $34.

JARVIS, M.G., HAZELDEN, J. and MACKNEY, D. Soils of Berkshire. Soil Survey Bulletin No.8. Pp.263. (Soil Survey of England and Wales, Rothamsted Experimental Station, Harpenden, Herts: 1979.) Np.

JARVIS, M.G. and MACKNEY, D. (eds.) Soil Survey Applications. Technical Monograph No.13. Pp.197. (Soil Survey of England and Wales, Rothamsted Experimental Station, Harpenden, Herts: 1979.) Flexi £4.

KENNEDY, V. S. (ed.). Estuarine Perspectives. Proceedings of the 5th Biennial Research Conference, Jekyll Island, Georgia, October 1979. Pp.533. ISBN 0-12-404060-8. (Academic: 1980.) $29.50.

KENT, P. Minerals from the Marine Environment. Pp.88. Flexi ISBN 0-7131-2813-5. (Edward Arnold: 1981.) £3.95.

KIRKBY, M. J. (ed.). Soil Erosion. Pp.312. ISBN 0-471-27802-5. (Wiley: 1981.) £22.50, $67.

LANDING, E., LUDVIGSEN, R. and von BITTER, P.H. Upper Cambrian to Lower Ordovician Conodont Biostratigraphy and Biofacies, Rabbitkettle Formation, District of Mackenzie. Life Sciences Contributions Royal Ontario Museum, Number 126. Pp.42. Pbk ISBN 0-88854-265-8. (Royal Ontario Museum, Toronto: 1980.) $3.

LAVILLE, H., RIGAUD, J.-P. and SACKETT, J. Rock Shelters of the Perigord: Geological Stratigraphy and Archaeological Succession. Pp.371. ISBN 0-12-438750-0. (Academic: 1980.) $29.50.

SHUVAL, H.I. Water Quality Management under Conditions of Scarcity. Israel as a Case Study. Pp.352. ISBN 0-12-641280-4. (Academic: 1980.) $38.

STUCKI, F. W. and BANWART, W. L. (eds). Advanced Chemical Methods for Soil and Clay Minerals Research. Proceedings of the NATO Advanced Study Institute held at the University of Illinois, July 1979. Pp.477. ISBN 90-277-1158-5. (Reidel: 1980.) Dfl. 110, $58.

TCHERNIA, P. Descriptive Regional Oceanography: An Elementary Description of the Four Main Divisions of the World Ocean, of their Limits, Forms, Topography, Wind Systems, Climatology, Surface Circulation and Hydrological Characteristics and Structure. Pergamon Marine Series, Vol.3. Pp.253. Hbk ISBN 0-08-020925-4; flexi ISBN 0-08-020919-X. (Pergamon: 1980.) Hbk $40, £20; flexi $17.50, £8.

Biological Sciences

AGARWAL, M.K. (ed.). Bacterial Endotoxins and Host Response. Proceedings of the 4th International Congress of Immunology Satellite Workshop held in Paris, July 1980. Pp.436. ISBN 0-444-80301-7. (Elsevier/North-Holland Biomedical: 1980.) $65, Dfl. 133.

ARBER, W. *et al.* (eds). Current Topics in Microbiology and Immunology, Vol.90. Pp. 147. ISBN 3-540-10181-0. (Springer: 1980.) DM 88, $52.

BARNES, H. and BARNES, M. (eds). Oceanography and Marine Biology: An Annual Review, Vol.18. Pp.528. ISBN 0-08-025732-1. (Pergamon: 1980.) £35, $84.

BEAN, W.J. Trees and Shrubs Hardy in the British Isles. 8th Edn, Revised. Vol. IV Ri–Z. Pp.808. ISBN 0-7195-2428-8. (John Murray: 1980.) £40.

BEARMAN, D. and EDSALL, J. T. (eds). Archival Sources for the History of Biochemistry and Molecular Biology: A Reference Guide and Report. Pp.338. ISBN 0-8357–0511-0. (American Philosophical Society, Philadelphia: 1980.) $28.

BOSSCHE, H. Van den. (ed.). The Host–Invader Interplay. Janssen Research Foundation Series, Vol.2. Proceedings of the 3rd International Symposium on the Biochemistry of Parasites and Host–Parasite Relationships, Beerse, Belgium, June 1980. Pp.766. ISBN 0-444-80284-3. (Elsevier/North-Holland Biomedical: 1980.) $83, Dfl. 170.

BOURNE, G. H. (ed.). Hearts and Heart-Like Organs. Vol. 1, Comparative Anatomy and Development. Pp.415. ISBN 0-12-119401-9. (Academic: 1980.) $51.

BOURNE, G. H. (ed.). Hearts and Heart-Like Organs. Vol.2, Physiology. Pp.573. ISBN 0-12-119402-7. (Academic: 1980.) $57.

BOURNE, G. H. and DANIELLI, J. F. (eds). International Review of Cytology, Vol.68. Pp.327. ISBN 0-12-364468-2. (Academic: 1980.) $38.

CATTON, W. R. Jr. Overshoot: The Ecological Basis of Revolutionary Change. Pp.298. ISBN 0-252-00818-9. (University of Illinois Press: 1981.) £6.60, $9.90.

CHATT, J., da CÂMARA PINA, GENERAL L. M. and RICHARDS, R. L. (eds). New Trends in the Chemistry of Nitrogen Fixation. Pp.284. ISBN 0-12-169450-X. (Academic: 1980.) £22.40, $52.

DICKINSON, A. Contemporary Animal Learning Theory. Pp.177. Hbk ISBN 0-521-23469-7; pbk ISBN 0-521-29962-4. (Cambridge University Press: 1981.) Hbk £12.50; pbk. £3.95.

DIXON, F. J. and KUNKEL, H. G. (eds). Advances in Immunology, Vol.30. Pp.344. ISBN 0-12-022430-5. (Academic: 1980.) $35.

DOKULIL, M., METZ, H. and JEWSON, D. (eds). Shallow Lakes — Contributions to their Limnology. Developments in Hydrobiology, Vol.3. Proceedings of a Symposium held at Illmitz, Austria, September 1979. Pp.218. ISBN 90–6193-753-1. (W. Junk, The Hague: 1980.) Dfl. 115,$59.50.

ELLWOOD, D. C. *et al.* (eds). Contemporary Microbial Ecology. Proceedings of the 2nd International Symposium, University of Warwick, September 1980. Pp.438. ISBN 0-12-236550-X. (Academic: 1980.) £18, $41.

HICKEY, M. and KING, C. 100 Families of Flowering Plants. Pp.567. Hbk ISBN 0-521-23283-X; pbk ISBN 0-521-29891-1. (Cambridge University Press: 1981.) Hbk £27.50; pbk £8.95.

HORECKER, B. L. and STADTMAN, E. R. (eds). Current Topics in Cellular Regulation, Vol.17. Pp.308. ISBN 0-12-152817-0. (Academic: 1980.) $35.

Applied Biological Sciences

NORPOTH, K.H. and GARNER, R.C. (eds). Short-Term Test Systems for Detecting Carcinogens. Pp.417. ISBN 3-540-09203-X. (Springer-Verlag: 1980.) DM 72, $40.40.

NUDELMAN, S. and PATTON, D. D. (eds). Imaging for Medicine. Vol.1. Nuclear Medicine, Ultrasonics and Thermography. Pp.498. ISBN 0-306-40384-6. (Plenum: 1980.) $49.50.

PARK, W.W. The Histology of Borderline Cancer (With Notes on Prognosis). Pp.471. ISBN 3-540-09792-9. (Springer-Verlag: 1980.) DM 125, $73.80.

PINEDO, H. M. (ed.). Cancer Chemotherapy 1980. The EORTC Cancer Chemotherapy Annual 2. Pp.486. ISBN 90-219-3054-4. (Excerpta Medica: 1980.) $78, Dfl.160.

RAMWELL, P. (ed.). Prostaglandin Synthetase Inhibitors: New Clinical Applications. Prostaglandins and Related Lipids, Vol.1. Proceedings of the Symposium Les Prostaglandins en Clinique Humaine, Paris, September 1979. Pp.410. ISBN 0-8451-2100-6. (Alan R. Liss: 1980.) $34, Dfl. 102.

REINERT, J. (ed.). Chloroplasts. Results and Problems in Cell Differentiation, Vol.10. Pp.240. ISBN 3-540-10082-2. (Springer-Verlag: 1980.) DM 78, $46.10.

RINNE, U.K., KLINGER, M. and STAMM, G. (eds). Parkinson's Disease. Current Progress, Problems and Management. Proceedings of the Northern European Symposium held in Helsinki, November 1979. Pp.401. ISBN 0-444-80263-0. (Elsevier/North-Holland: 1980.) $65, Dfl.133.

SANADI, D.R. Current Topics in Bioenergetics, Vol.10. Pp.297. ISBN 0-12-152510-4. (Academic: 1980.) $34.50.

SCHERER, E. (Herausgeber). Strahlentherapie. Radiologische Onkologie. Pp.1003. ISBN 3-540-09780-5. (Springer-Verlag: 1980.) DM 198, $116.90.

SCHWARTZ, D., FLAMANT, R. and LELLOUCH, J. Clinical Trials. Flexi ISBN 0-12-632440-9. (Academic: 1980.) £15, $34.50.

SELLERS, A.H. UICC Technical Report Series. Vol.51, TNM Classification of Malignant Tumours: A Brochure of Checklists. Pp.50. Loose-leaf ISBN 92-9018-051-X. (International Union Against Cancer, Geneva: 1980.) SwFr.20.

SILLIKER, J.H. (chairman, ed. committee). Microbial Ecology of Foods. Vol.II, Food Commodities. International Commission on Microbiological Specifications for Foods. Microorganisms in Foods 3. Pp.997. ISBN 0-12-363502-0. (Academic: 1980.) $60.

SONNENWIRTH, A.C. and JARETT, L. (eds). Gardwohl's Clinical Laboratory Methods and Diagnosis, Vol.1, 8th Edn. Pp.1206 + index; Vol.2, 8th Edn. Pp.1130 + index. ISBN 0-8016-4741-X. (YB Medical, London: 1980.) £85 (2 vol. set).

TREVAN, M. D. Immobilized Enzymes: An Introduction and Applications in Biotechnology. Pp.138. ISBN 0-471-27826-2. (Wiley: 1980.) £8.75.

TURNER, P. (ed.). Clinical Pharmacology and Therapeutics. Proceedings of Plenary Lectures, Symposia and Therapeutic Sessions of the 1st World Conference on Clinical Pharmacology & Therapeutics, London, August 1980. Pp.576. Hbk ISBN 0-333-31229-5; pbk ISBN 0-333-31230-9. (Macmillan, London: 1980.) Np.

VAN DER MAAREL, E., ORLOCI, L. and PIGNATTI, S. (eds). Data-processing in Phytosociology. Advances in Vegetation Science, 1. Report on the Activities of the Working-group of the International Society for Vegetation Science, 1969–1978. Pp. 228. ISBN 90-6193-608-X. (W. Junk, The Hague: 1980.) Dfl. 150, $79.

VAN VUNAKIS, H. and LANGONE, J.J. (eds). Methods in Enzymology, Vol.70. Immunochemical Techniques Part A. Pp.525. ISBN 0-12-181879-1. (Academic: 1980.) $48.

VERHOEF, J., PETERSON, P.K. and QUIE, P.G. (eds). Infections in the Immunocompromised Host—Pathogenesis, Prevention and Therapy. Developments in Immunology, Vol.11. Proceedings of the First International Symposium held in Veldhoven, The Netherlands, June 1980. Pp.315. ISBN 0-444-80287-8. (Elsevier/North-Holland: 1980.) $39, Dfl.80.

VILLAZON, A.S. et al. (eds). Critical Care Medicine. International Congress Series 503. Proceedings of the First Panamerican Congress held in Mexico City, September 1979. Pp.230. ISBN 90-219-0431-4. (Excerpta Medica: 1980.) $53.75, Dfl.110.

WORF, D.L. (ed.). Biological Monitoring for Environmental Effects. Pp.227. ISBN 0-669-03306-5. (Gower Press, Farnborough, Hampshire, UK: 1980.) £15.50.

WORLD HEALTH ORGANIZATION. Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal. IARC Monographs, Supplement 2. Reports of IARC *ad hoc* Working Groups, Hanover, FRG, June 1979. Pp.426. Flexi ISBN 92-832-1405-6. (International Agency for Research on Cancer, Lyon: 1980.) Sw.Fr. 40, $25.

Psychology

ANDERSON, J.R. and BOWER, G.H. Human Associative Memory: A Brief Edition. Pp.260. Flexi ISBN 0-89859-020-5. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $14.95.

BELL, R.W. and SMOTHERMAN, W.P. (eds). Maternal Influences and Early Behavior. Pp.442. ISBN 0-85200-537-7. (MTP, Lancaster: 1980.) £24.95.

BOWER, G.H. (ed.) The Psychology of Learning and Motivation: Advances in Research and Theory. Pp.367. ISBN 0-12-543314-X. (Academic: 1980.) $36.

BURROWS, G.D. and DAVIES, B. (eds.) Handbook of Studies on Anxiety. Pp.405. ISBN 0-444-80224-X. (Elsevier/North-Holland: 1980.) $79, Dfl.162.

COLLINS, W.A. (ed.). Development of Cognition, Affect, and Social Relations. The Minnesota Symposia on Child Psychology, Vol.13. Pp.294. ISBN 0-89859-023-X. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $19.95.

FRENCH, P. and MACLURE, M. (eds). Adult–Child Conversation. Pp.310. ISBN 0-7099-0069-4. (Croom Helm: 1981.) £14.95.

GREGG, L.W. and STEINBERG, E.R. (eds). Cognitive Processes in Writing. Pp.177. ISBN 0-89859-032-9. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $19.95.

HEMMINGS, G. (ed.). Biochemistry of Schizophrenia and Addiction: In Search of a Common Factor. Pp.344. ISBN 0-85200-310-2. (MTP: Lancaster, 1980.) £17.95.

HONECK, R.P. and HOFFMAN, R.R. (eds). Cognition and Figurative Language. Pp.436. ISBN 0-89859-047-7. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $29.95.

JEEVES, M.A. (ed.). Psychology Survey No.3. Pp.199. Hbk ISBN 0-04-150073-3; pbk ISBN 0-04-150074-1. (Allen and Unwin: London, 1980.) Hbk £10.50; pbk £4.95.

KIDD, R.F. and SAKS, M.J. (eds). Advances in Applied Social Psychology, Vol.1. Pp.222. ISBN 0-89859-027-2. (Lawrence Erlbaum Associates, Hillsdale, New Jersey: 1980.) $19.95.

KIRBY, J.R. and BIGGS, J.B. (eds). Cognition, Development, and Instruction. Pp.217. ISBN 0-12-409550-X. (Academic: 1980.) $19.

KLUWE, R.H. and SPADA, H. (eds). Developmental Models of Thinking. Pp.304. ISBN 0-12-416450-1. (Academic: 1980.) $26.50.

RATHJEN, D. P. and FOREYT, J. P. Social Competence: Interventions for Children and Adults. Published in Cooperation with the Houston Behaviour Therapy Association. Pergamon General Psychology Series, Vol.91. Pp.337. ISBN 0-08-025965-0. (Pergamon: 1980.) £13.25.

SARASON, I. G. and SPIELBERGER, C. D. (eds). Stress and Anxiety, Vol.7. Pp.308. ISBN 0-89116-183-X. (Hemisphere Publishing: 1980.) $25.50.

Sociology

GEBER, B.A. and NEWMAN, S.P. Soweto's Children: The Development of Attitudes. European Monographs in Social Psychology 20. Pp.215. ISBN 0-12-278750-1. (Academic: 1980.) £13.60, $41.50.

WILLIAMS, W. Government by Agency. Lessons from the Social Program Grants-in-Aid Experience. Pp.289. ISBN 0-12-755950-7. (Academic: 1980.) $22.50.

Anthropology

BARLETT, P.F. (ed.). Agricultural Decision Making. Anthropological Contributions to Rural Development. Pp.378. ISBN 0-12-078880-2. (Academic: 1980.) $26.

COCKBURN, A. and E. (eds). Mummies, Disease, and Ancient Cultures. Pp.340. ISBN 0-521-23020-9. (Cambridge University Press: 1980.) £25.

General

ARTHURS, A. M. Complementary Variational Principles. 2nd Edn. Pp.154. ISBN 0-19-853532-5. (Clarendon/Oxford University Press: 1980.) £16.

BENARIE, M. M. Urban Air Pollution Modelling. Pp.405. ISBN 0-262-02140-4. (MIT Press: 1980.) $45.

BERG, C. P. The University of Iowa and Biochemistry from their beginnings. Pp.173. ISBN 0-87414-014-5. (University of Iowa: 1980.) Np.

BROOK, P. J. (ed.). Natural History of Auckland: An Introduction. Pp.80. (Auckland War Memorial Museum: 1979.) Pbk Np.

CHABALLE, L. Y., MASUY, L. and VANDENBERGHE, J. -P. (Compiled by.) Elsevier's Oil and Gas Field Dictionary. In six languages: English/American, French, Spanish, Italian, Dutch and German. Pp.674. ISBN 0-444-41833-4. (Elsevier Scientific: 1980.) $139, Dfl.285.

CLARK, R. W. The Greatest Power on Earth, The Story of Nuclear Fission. Pp.342. ISBN 0-283-98715-4. (Sidgwick & Jackson: London, 1980.) £8.95.

DAVIDGE, R. and WOODING, E. R. The Use of Scientific Information Services. 4th Edn. Pp.50. (Royal Holloway College, Surrey, 1980.) £2.

DEVORE, R.A. and SCHERER, K. Quantative Approximation. Proceedings of a Symposium held in Bonn, August 1979. Pp.324. ISBN 0-12-213650-0. (Academic: 1980.) $22.

GARLAND, M.M. Cambridge Before Darwin. The Ideal of a Liberal Education 1800-1860. Pp.196. ISBN 0-521-23319-4. (Cambridge University Press: 1980.) £14.50.

GIARINI, O. Dialogue on Wealth and Welfare. An Alternative View of World Capital Formation. A Report to the Club of Rome. Pp.386. Hbk ISBN 0-08-026088-8; flexi ISBN 0-08-026087-X. (Pergamon: 1980.) Hbk £20, $47; flexi £10, $23.50.

GOULD, D. J. Bureaucratic Corruption and Underdevelopment in the Third World: The Case of Zaire. Pp.181. ISBN 0-08-025084-X. (Pergamon: 1980.) £11.25.

HALL, T.E., JONES, P.R. and PRESTON, G.B. (eds). Semigroups. Proceedings of the Monash University Conference, Monash University, Victoria, Australia, October 1979. Pp.255. ISBN 0-12-319450-4. (Academic: 1980.) $18.

HARRIS, J. E. and WENTE, E. F. (eds). An X-Ray Atlas of the Royal Mummies. Pp.403. ISBN 0-226-31745-5. (University of Chicago Press: London.) £33.

LEFER, A. M., SEBA, T. M. and MELA, L. M. (eds). Advances in Shock Research, Vol.3. Papers Presented at the 2nd Annual Conference, Williamsburg, Virginia, June 1979. Pp.296. ISBN 0-8451-0602-3. (Alan R. Liss: New York, 1980.) $34.

MILLER, R.G. Jr. et al. (eds). Biostatistics Casebook. Pp.238. Flexi ISBN 0-471-06258-8. (Wiley: 1980.) £7.95.

MYERS, P. A. An Introduction to Radiographic Technique. Pp.115. Flexi (spiral binding) ISBN 0-03-056654-1. (Praeger Publishers CBS Educational and Professional Publishing: New York, 1980.) Np.

NELKIN, D. and POLLAK, M. Atom Besieged. Extra Parliamentary Dissent in France and Germany. Pp.235. ISBN 0-262-14034-9. (MIT Press: 1980.) $17.50.

OPEC. Opec and Future Energy Markets. Proceedings of the OPEC Seminar, Vienna, Austria, October 1979. Prepared by the Public Information Department of Opec. Pp.290. ISBN 0-333-31126-4. (Macmillan: 1980.) £20.

PASK, G., GLANVILLE, R. and ROBINSON, R. Calculator Saturnalia or Travels with a Calculator. A Compendium of Diversions and Improving Exercises for Ladies and Gentlemen. Pp.256. ISBN 0-7045-3007-4. (Wildwood House: London, 1980.) £7.95.

PENTREATH, R. J. Nuclear Power, Man and the Environment. Pp.255. Flexi ISBN 0-85109-840-1. (Taylor & Francis: London, 1980.) £6.

PICHLER, F. R. and HANIKA, F. de P. (eds). Progress in Cybernetics and Systems Research. Vol.VII: General Systems Methodology Organization and Management Cognition and Learning. Pp.393. ISBN 0-89116-195-3. (Hemisphere Publishing: 1980.) $50.

PIERCE, F. R. and POSNER, E. C. Introduction to Communication Science and Systems. Pp.390. ISBN 0-306-40492-3. (Plenum: 1980.) $27.50.

PLAGE, D. Wild Horizons. Pp.216. ISBN 0-00-216029-3. (Collins: 1980.) £8.95.

ROSSEAU, G. S. and PORTER, R. (eds). The Ferment of Knowledge. Pp.500. ISBN 0-521-22599-X. (Cambridge University Press: 1980.) Np.

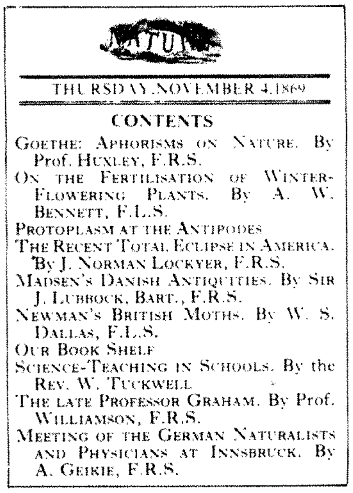
SCHUMER, W., SPITZER, J. J. and MARSHALL, B. E. (eds). Advances in Shock Research, Vol.4. Papers Presented at the 2nd Annual Conference, Williamsburg, Virginia, June 1979. Pp.210. ISBN 0-8451-0603-1. (Alan R. Liss: New York, 1980.) $26.

SEYMOUR, I. OPEC Instrument of Change. Pp.306. ISBN 0-333-30667-8. (Macmillan: 1980.) £15.

Fascinating reading and an unusual gift

A high-quality, 40-page facsimile of the first issue of **Nature** is now available at the following prices:

UK	**75p**
Rest of World	**US $1.50** (Surface)
Rest of World	**US $2.00** (Air)



THURSDAY, NOVEMBER 4, 1869

CONTENTS

Goethe: Aphorisms on Nature. By Prof. Huxley, F.R.S.
On the Fertilisation of Winter-Flowering Plants. By A. W. Bennett, F.L.S.
Protoplasm at the Antipodes
The Recent Total Eclipse in America. By J. Norman Lockyer, F.R.S.
Madsen's Danish Antiquities. By Sir J. Lubbock, Bart., F.R.S.
Newman's British Moths. By W. S. Dallas, F.L.S.
Our Book Shelf
Science-Teaching in Schools. By the Rev. W. Tuckwell
The late Professor Graham. By Prof. Williamson, F.R.S.
Meeting of the German Naturalists and Physicians at Innsbruck. By A. Geikie, F.R.S.

Payment may be made in any currency at the prevailing exchange rate. Cheques should be made payable to **Nature.** Prices include postage and packing.

To order please complete the coupon below, or write, enclosing the correct remittance.

Order Form

To: Nature
Macmillan Journals Ltd
Brunel Road, Basingstoke
Hants, RG21 2XS, England

Please send me .... facsimile(s) of the first issue of NATURE.

I enclose..

Name ....

Address ....

....

....

....

Reg. No: 785998 England.
Reg. Office: 4 Little Essex St, London WC2R 3LF

Circle No. 56 on Reader Enquiry Card.

SHORR, J. E., SOBEL, G. E., ROBIN, P. and CONNELLA, J. A. (eds). Imagery; Its Many Dimensions and Applications. Proceedings of the First Annua Conference of the American Association for the Study of Mental Imagery, held at Los Angeles, June 1979. Pp.398. ISBN 0-306-4056-7. (Plenum: 1980.) $32.50.

SKUTCH, A. F. A. Naturalist on a Tropical Farm. Pp.397. ISBN 0-520-03802-9. (University of California Press: 1980.) £10.25.

SMIL, V. and KNOWLAND, W. E. (eds). Energy in the Developing World. The real energy crisis. Pp.386. Hbk ISBN 0-19-854425-1; pbk ISBN 0-19-854421-9. (Oxford University Press: 1980.) Hbk £25; pbk £12.50.

STIBIC, V. Personal Documentation for Professionals. Means and Methods. Pp.214. ISBN 0-444-85480-0. (North-Holland: 1980.) $29.25, Dfl. 60.

WHEELWRIGHT, H. W. Bush Wanderings of a Naturalist. Pp.272. ISBN 0-19-54226-6. (Oxford Unversity Press: 1980.) £8.95.

WOOD, C. (ed.). Human Health and Environmental Toxicants. Royal Society of Medicine International Congress and Symposium Series No. 17. Pp.210. Flexi ISBN 0-12-794919-4. (Academic: 1980.) £15, $36.

YALLOP, H. J. Explosion Investigation. Pp.276. ISBN 0-7073-0272-2. (Scottish Academic Press: 1980.) £20.

Mathematics

ANTOINE, J.-P. and TIRAPEGUI, E. (eds). Functional Integration Theory and Applications. Proceedings of the Colloquium held in Louvain-la-Neuve, Belgium, November 1979. Pp.355. ISBN 0-306-40573-3. (Plenum: 1980.) $42.50.

CLEMENS, C.H. A Scrapbook of Complex Curve Theory. Pp.186. ISBN 0-306-40536-9. (Plenum: 1980.) $22.50.

FRAUENTHAL, J.C. Mathematical Modeling in Epidemiology. Pp.118. Flexi ISBN 3-540-10328-7. (Springer-Verlag: 1980.) DM 36, $21.30.

MISCHENKO, E.F. and ROZOV, N.K. Differential Equations with Small Parameters and Relaxation Oscillations. Mathematical Concepts and Methods in Science and Engineering, Vol.13. Pp.228. ISBN 0-306-39253-4. (Plenum: 1980.) $29.50.

PLUMPTON, C. and MacILWAINE, P.S.W. New Tertiary Mathematics. Vol.1, Part 1, Pure Mathematics: The Core. Pp.401. Hbk ISBN 0-08-025031-9. (Pergamon: 1980.) Hbk £17.40; $40, flexi £6.90, $15.95.

PLUMPTON, C. and MacILWAINE, P.S.W. New Tertiary Mathematics. Vol.1, Part 2, Basic Applied Mathematics. Pp.229. Hbk ISBN 0-08-025035-1; flexi ISBN 0-08-021645-5. (Pergamon: 1980.) Hbk £17.40, $40; flexi £6.90, $15.95.

PLUMPTON, C. and MacILWAINE, P.S.W. New Tertiary Mathematics. Vol.2, Part 1, Further Pure Mathematics. Pp.803. Hbk ISBN 0-08-025033-5; flexi ISBN 0-08-021644-7. (Pergamon: 1980.) Hbk £17.40, $40; flexi £6.90, $15.95.

MARCHUK, G.I. and NISEVICH, N.I. (eds). Mathematical Methods in Clinica Practice. Pp.116. ISBN 0-08-025493-4. (Pergamon: 1980.) £25, $60.

TEVES, F. Introduction to Pseudodifferential and Fourier Integral Operator Vol.1, Pseudodifferential Operators. Pp.299. ISBN 0-306-40403-6. (Plenum: 1980.) Np.

WOODHOUSE, N. Geometric Quantization. Pp.316. ISBN 0-19-853528-7. (Clarendon Press/Oxford University Press: 1981.) £27.50.

Astronomy

HEIDMANN, J. Relativistic Cosmology. An Introduction. Pp.168. Flexi ISBN 3-540-10138-1. (Springer-Verlag: 1980.) DM 48, $28.30.

WAYMAN, P.A. (ed.). Proceedings of the 17th General Assembly, Montreal 1979. Transactions of the International Astronomical Union, Vol.XVIIB. Pp.526. ISBN 90-277-1159-3. (Reidel: 1980.) Dfl.130, $68.50.

Physics

BALLHAUSEN, C.J. and GRAY, H.B. Molecular Electronic Structures: An Introduction. Pp.133. ISBN 0-8053-0452-5. (Benjamin/Cummings, California: 1980.) $14.50.

BALTES, H.P. (ed.). Inverse Scattering Problems in Optics. Topics in Current Physics, Vol.20. Pp.313. ISBN 3-540-10104-7. (Springer-Verlag: 1980.) DM 78, $46.10.

BECK, R., ENGLISH, W. and GÜRS, K. Table of Laser Lines in Gases and Vapors. 3rd Edn. Springer Series in Optical Sciences, Vol.3. Pp.247. ISBN 3-540-10347-3. (Springer-Verlag: 1980.) DM 62, $36.50.

BELLAMY, L.J. The Infrared Spectra of Complex Molecules. Vol.2, Advance in Infrared Group Frequencies. 2nd Edn. Pp.299. ISBN 0-412-22350-3. (Chapman & Hall: 1980.) £15.

BOLTON, W. Fields. Study Topics in Physics, Book 5. Pp.95. Flexi ISB 0-408-10656-5. (Butterworths: 1980.) £1.60.

BOLTON, W. Rays, Waves and Oscillations. Study Topics in Physics, Boo Pp.95. Flexi ISBN 0-408-10657-3.

BRUCE, A.D. and COWLEY, R.A. Structural Phase Transitions. Pp.326. IS 0-85066-206-0. (Taylor & Francis, London: 1981.) £15.

CANNATA, F. and ÜBERALL, H. Giant Resonance Phenomena Intermediate-Energy Nuclear Reactions. Springer Tracts in Modern Physic Vol.89. Pp.112. ISBN 3-540-10105-5. (Springer-Verlag: 1980.) DM 52, $30.60.

CUNNINGHAM, R.E. and WILLIAMS, R.J.J. Diffusion in Gases and Poro Media. Pp.275. ISBN 0-306-40537-7. (Plenum: 1980.) Np.

DEGIORGIO, V., CORTI, M. and GIGLIO, M. (eds). Light Scattering Liquids and Macromolecular Solutions. Proceedings of the Workshop, Cen Informazioni Studi Esperienze, Segrate, Milan, May 1979. Pp.295. IS 0-306-40558-X. (Plenum: 1980.) $35.